

Copyright (c) 1993 - 2004 Compugen Ltd.

Gencore version 5.1.6

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:15:59 ; Search time 20.1923 Seconds
(without alignments)

Scoring table: BloSUM62

Score: Gapext 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%
Listing First 45 summaries

Database :	Issued_Patents_AA.*
	1: /cgm2_6/podata/2/iaa/5A_COMB.pep:*
	2: /cgm2_6/podata/2/iaa/5B_COMB.pep:*
	3: /cgm2_6/podata/2/iaa/6A_COMB.pep:*
	4: /cgm2_6/podata/2/iaa/6B_COMB.pep:*
	5: /cgm2_6/podata/2/iaa/SPCTUS_COMB.pep:*
	6: /cgm2_6/podata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	142	100.0	25	4	US-09-314-268-4
2	121	85.2	92	4	US-09-314-268-90
3	110.5	77.8	30	2	US-08-334-915-40
4	98.5	69.4	20	2	US-08-334-915-40
5	98	69.0	35	4	US-08-314-268-117
6	93	65.5	20	2	US-08-314-268-117
7	91	64.1	15	4	US-09-314-268-167
8	75	52.8	12	4	US-09-314-268-168
9	59.5	41.9	39	4	US-09-314-268-119
10	57	40.1	34	4	US-09-314-268-118
11	56	39.8	38	4	US-09-314-268-120
12	56.5	39.8	49	4	US-09-314-268-93
13	56	39.4	31	4	US-09-314-268-114
14	55.6	39.4	32	4	US-09-314-268-114
15	55.5	39.1	442	4	US-09-314-268-114
16	54	38.0	173	4	US-09-314-268-113
17	53	37.3	143	4	US-09-314-268-113
18	53	37.3	202	4	US-09-314-268-114
19	52.5	37.0	182	4	US-09-314-268-114
20	52	36.6	8	4	US-09-314-268-31
21	52	36.6	8	4	US-09-314-268-33
22	52	36.6	8	4	US-09-314-268-37
23	52	36.6	184	4	US-09-314-268-11
24	51.5	36.3	117	4	US-09-314-268-92
25	51	35.9	8	4	US-09-314-268-36
26	51	35.9	201	3	US-08-987-418A-2
27	51	35.9	201	3	US-09-343-062-2

Patent No. 5498499
Sequence 9668, App
Sequence 23, App
Sequence 23, App
Sequence 2, App
Sequence 2, App
Sequence 7, App
Patent No. 5196511
Sequence 34, App
Sequence 35, App
Sequence 38, App
Sequence 12273, A
Sequence 18484, A
Sequence 19129, A
Sequence 22162, A
Sequence 25003, A
Sequence 105, App
Sequence 32, App

ALIGMENTS

RESULT 1 US-09-314-268-4	RESULT 2 US-09-314-268-90
; Sequence 4, Application US/09314268	; Sequence 4, Application US/09314268
; Patent No. 6346377	; Patent No. 6346377
; GENERAL INFORMATION:	; GENERAL INFORMATION:
; APPLICANT: Doobhar, John	; APPLICANT: Doobhar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA	; TITLE OF INVENTION: VIRUSS
; FILE REFERENCE: 3789/80902	; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314-268	; CURRENT APPLICATION NUMBER: US/09/314-268
; CURRENT FILING DATE: 1999-03-19	; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314-268	; EARLIER APPLICATION NUMBER: 09/314-268
; NUMBER OF SEQ ID NOS: 179	; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentin Ver. 2.1	; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 4	; SEQ ID NO: 4
; LENGTH: 25	; LENGTH: 25
; TYPE: PRT	; TYPE: PRT
; ORGANISM: Human papillomavirus type 16	; ORGANISM: Human papillomavirus type 16
US-09-314-268-4	US-09-314-268-4

Best Local Similarity 91.3%; Pred. No. 5.5e-07; Mismatches 0; Indels 0; Gaps 2; SEQ ID NO 167

RESULT 6
US-08-934-915-170
; Sequence 170 Application US/08934915
; Patent No. 5332412

; GENERAL INFORMATION:
; APPLICANT: DILINER, LENA
; APPLICANT: CHENG, HWEI-MING
; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN PAPILLOMAVIRUS 1, 5, 6, 8, 11, 16, 18, 31, 33 AND 56, USEFUL IN IMMUNOASSAY FOR DIAGNOSTIC PURPOSES
; NUMBER OF SEQUENCES: 193
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MASON & ASSOCIATES, P.A.
; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
; CITY: CLEARWATER
; STATE: FLORIDA
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 3.0
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/934,915
; FILING DATE: 22-SEP-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/949,836
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: LOUISE A. FOUTCH
; REGISTRATION NUMBER: 37,133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 813-518-3800
; TELEFAX: 813-538-3820
; INFORMATION FOR SEQ ID NO: 170:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-934-915-170

; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314,268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314,268
; EARLIER FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 167
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-314-268-167

Query Match 64.1%; Score 91; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPIPKSPWPKKHKR 15
Db 1 RPIPKSPWPKKHKR 15

RESULT 8
US-09-314-268-168
; Sequence 168 Application US/09314268
; Patent No. 6346377

; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; TITLE OF INVENTION: VIRUSES
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314,268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314,268
; EARLIER FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 168
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-314-268-168

Query Match 52.8%; Score 75; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PKPSWPKKHKR 15
Db 1 PKPSWPKKHKR 12

RESULT 9
US-09-314-268-119
; Sequence 119 Application US/09314268
; Patent No. 6346377

; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; TITLE OF INVENTION: VIRUSES
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314,268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314,268
; EARLIER FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 119
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Human papillomavirus type 31
US-09-314-268-119

Query Match 65.5%; Score 93; DB 2; Length 20;
Best Local Similarity 80.0%; Pred. No. 1.5e-06;
Matches 16; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 IPKPSWPKKHKRLSDQDS 22
Db 1 IPKPSWPKKHKRISQN 20

RESULT 7
US-09-314-268-167
; Sequence 167 Application US/09314268
; Patent No. 6346377

; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; TITLE OF INVENTION: VIRUSES


```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 114
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Human papillomavirus type 59
; US-09-314-268-114

Query Match      39.4%; Score 56; DB 4; Length 32;
Best Local Similarity 63.6%; Pred. No. 0.3; Mismatches 3; Indels 4; Gaps 2;
Matches 14; Conservative 1; Missmatches 3; Indels 4; Gaps 2;

Qy      5 KPSPWAPKK--HRRL-SDQDS 22
Db       1 KPRTWAPKRGTVRRRLESDDQS 22

RESULT 15
US-09-252-991A-31743
; Sequence 31743, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOVIRUSES FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO: 31743
; LENGTH: 442
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-31743

Query Match      39.1%; Score 55.5; DB 4; Length 442;
Best Local Similarity 48.0%; Pred. No. 6.2; Mismatches 11; Indels 1; Gaps 1;
Matches 12; Conservative 1; Missmatches 11; Indels 1; Gaps 1;

Qy      1 RPIPKSPWAPKCHRRLSDQDSQTP 25
Db       1 RPFPFHGPQFRHR-GDQGRQVP 161

Search completed: May 27 2004 16:22:26

```

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:21:04 ; Search time 52.8846 Seconds

(Without alignments)
 132.009 Million cell updates/sec

Title: US-10-008-524A-4
 Perfect score: 142
 Sequence: 1.RPIPKSPWAPKRRRLSDQSQTP? 25

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Published Applications AA:
 1: /cgns2_6/ptodata/2/pubpaas/us07_pubcomb.pep:*

2: /cgns2_6/ptodata/2/pubpaas/us07_pubcomb.pep:*

3: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

4: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

5: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

6: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

7: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

8: /cgns2_6/ptodata/2/pubpaas/us08_pubcomb.pep:*

9: /cgns2_6/ptodata/2/pubpaas/us09_pubcomb.pep:*

10: /cgns2_6/ptodata/2/pubpaas/us09b_pubcomb.pep:*

11: /cgns2_6/ptodata/2/pubpaas/us09c_pubcomb.pep:*

12: /cgns2_6/ptodata/2/pubpaas/us09_new_pub.pep:*

13: /cgns2_6/ptodata/2/pubpaas/us1a_pubcomb.pep:*

14: /cgns2_6/ptodata/2/pubpaas/us1b_pubcomb.pep:*

15: /cgns2_6/ptodata/2/pubpaas/us10c_pubcomb.pep:*

16: /cgns2_6/ptodata/2/pubpaas/us10_new_pub.pep:*

17: /cgns2_6/ptodata/2/pubpaas/us60_pubcomb.pep:*

18: /cgns2_6/ptodata/2/pubpaas/us60_pubcomb.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match ID	Description
1	142	100.0	Sequence 5, Appli
2	131.5	92.6	Sequence 4, Appli
3	121	85.2	Sequence 90, Appli
4	121	85.2	Sequence 90, Appli
5	98	69.0	Sequence 117, Appli
6	98	69.0	Sequence 117, Appli
7	91	64.1	Sequence 167, Appli
8	82	57.7	Sequence 167, Appli
9	75	52.8	Sequence 168, Appli
10	75	52.8	Sequence 168, Appli
11	59.5	41.9	Sequence 119, Appli
12	59.5	41.9	Sequence 119, Appli
13	57	40.1	Sequence 118, Appli
14	57	40.1	Sequence 118, Appli
15	57	40.1	Sequence 152, Appli

%

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length	DB	Indels	Gaps	Organism
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16
Qy	1 RPIPKSPWAPKRRRLSDQSQTP	25					
Db	1 RPIPKSPWAPKRRRLSDQSQTP	25					
US-10-008-524A-4	100.0%	142,	25;	US-10-008-524A-4	0;	0;	Human papillomavirus type 16

RESULTS

Query Match ID	Best Local Similarity	Score	Length</th
----------------	-----------------------	-------	------------

FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; PRIOR APPLICATION NUMBER: 2003-01-24
; PRIOR APPLICATION DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB1/01176
; PRIOR APPLICATION DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-4

RESULT 3
US-10-008-524A-90
Query Match 92.6%; Score 131.5; DB 15; Length 26;
Best Local Similarity 96.2%; Pred. No. 2.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 RPIPKPSPWAPKXHRRU-SDQDSQTP 25
Db 1 RPIPKPSPWAPKXHRRUSSDQDSQTP 26

RESULT 3
US-10-008-524A-90
Query Match 92.6%; Score 131.5; DB 15; Length 26;
Best Local Similarity 96.2%; Pred. No. 2.2e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 RPIPKPSPWAPKXHRRU-SDQDSQTP 25
Db 1 RPIPKPSPWAPKXHRRUSSDQDSQTP 26

RESULT 5
US-10-008-524A-117
Query Match 85.2%; Score 121; DB 15; Length 92;
Best Local Similarity 92.6%; Pred. No. 1.4e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 RPIPKPSPWAPKXHRRU-SDQDSQTP 25
Db 26 RPIPKPSPWAPKXHRRUSSDQDSQTP 52

RESULT 6
US-10-350-719-117
Query Match 69.0%; Score 98; DB 14; Length 35;
Best Local Similarity 91.3%; Pred. No. 3.9e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 5 KPSWPWAPKXHRRU-SDQDSQTP 25
Db 1 KPSWPWAPKXHRRUSSDQDSQTP 23

RESULT 6
US-10-350-719-117
Query Match 69.0%; Score 98; DB 14; Length 35;
Best Local Similarity 91.3%; Pred. No. 3.9e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 5 KPSWPWAPKXHRRU-SDQDSQTP 25
Db 1 KPSWPWAPKXHRRUSSDQDSQTP 23

RESULT 4
US-10-350-719-90
Query Match 92.6%; Score 131.5; DB 15; Length 35;
Best Local Similarity 91.3%; Pred. No. 3.9e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 5 KPSWPWAPKXHRRU-SDQDSQTP 25
Db 1 KPSWPWAPKXHRRUSSDQDSQTP 23

FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; PRIOR APPLICATION NUMBER: US/10/350,719
; PRIOR APPLICATION DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB1/01176
; PRIOR APPLICATION DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 90
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-90

Query Match 85.2%; Score 121; DB 14; Length 92;
Best Local Similarity 92.6%; Pred. No. 1.4e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 1 RPIPKPSPWAPKXHRRU-SDQDSQTP 25
Db 26 RPIPKPSPWAPKXHRRUSSDQDSQTP 52

RESULT 4
US-10-350-719-90
Query Match 92.6%; Score 131.5; DB 15; Length 35;
Best Local Similarity 91.3%; Pred. No. 3.9e-05;
Matches 21; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

Qy 5 KPSWPWAPKXHRRU-SDQDSQTP 25
Db 1 KPSWPWAPKXHRRUSSDQDSQTP 23

RESULT 7
US-10-008-524A-167
; Sequence 167, Application US/10008524A
; Publication No. US20030175682A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 167
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-168

Query Match 52.8%; Score 75; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PKPSWAPKHR 15
Db 1 PKPSWAPKHR 12

RESULT 10
US-10-350-719-168
; Sequence 168, Application US/10350719
; Publication No. US20030219726A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 168
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-168

Query Match 52.8%; Score 75; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0099;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PKPSWAPKHR 15
Db 1 PKPSWAPKHR 12

RESULT 11
US-10-008-524A-119
; Sequence 119, Application US/10008524A
; Publication No. US20030175682A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 119
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Human papillomavirus type 31
US-10-008-524A-119

RESULT 9
US-10-008-524A-168
; Sequence 168, Application US/10008524A
; Publication No. US20030175682A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 168
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Human papillomavirus type 31
US-10-008-524A-119

Query Match 41.9%; Score 59.5; DB 14; Length 39;
 Best Local Similarity 46.7%; Pred. No. 2 4;
 Matches 14; Conservative 4; Mismatches 3; Indels 9; Gaps 2;

Db 5 KPSPWPKK----HRRLSDQD---SQTP 25
 Qy 1 KPAPWPKVCGGRRLSDQEQSQTETP 30

RESULT 12 US-10-350-719-119
 ; Sequence 119, Application US/10350719
 ; Publication No. US200301972A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Doorbar, John
 ; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
 ; FILE REFERENCE: 18396/2162
 ; CURRENT APPLICATION NUMBER: US/10/350-719
 ; CURRENT FILING DATE: 2003-01-24
 ; PRIOR APPLICATION NUMBER: PCT/GB01/01176
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: GB0018140.4
 ; PRIOR FILING DATE: 2000-07-24
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: PatentIn Ver. 2.1
 ; LENGTH: 34
 ; SEQ ID NO: 118
 ; TYPE: PRT
 ; ORGANISM: Human papillomavirus type 35
 ; US-10-350-719-118

Query Match 40.1%; Score 57; DB 15; Length 34;
 Best Local Similarity 58.8%; Pred. No. 4 3;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 5 KPSPWPKKHRRLSDQD 21
 Qy 1 KPAPWPKVCGGRRLSDQEQSQTETP 17

RESULT 15 US-10-080-170-152
 ; Sequence 152, Application US/10080170
 ; Publication No. US2003012960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: COLE, S.T.
 ; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR IDENTIFYING TARGETS FOR THE DIAGNOSTIS, PROPHYLAXIS OR TREATMENT OF MYCOBACTERIOSIS
 ; FILE REFERENCE: 03495.0218
 ; CURRENT APPLICATION NUMBER: US/10/080,170
 ; CURRENT FILING DATE: 2002-06-10
 ; PRIOR APPLICATION NUMBER: 60/270,123
 ; PRIOR FILING DATE: 2001-02-22
 ; NUMBER OF SEQ ID NOS: 652
 ; SEQ ID NO: 152
 ; SOFTWARE: PatentIn Ver. 2.1
 ; LENGTH: 230
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium leprae
 ; US-10-080-170-152

Query Match 40.1%; Score 57; DB 14; Length 230;
 Best Local Similarity 57.9%; Pred. No. 26;
 Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Db 4 PKSPWPKKHRRLSDQD 22
 Qy 110 PKPSKMAPRPLARLINGDS 128

Search completed: May 27, 2004, 16:34:18
 Job time : 53.8846 secs

Query Match 40.1%; Score 57; DB 14; Length 34;
 Best Local Similarity 58.8%; Pred. No. 4 3;
 Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 5 KPSPWPKKHRRLSDQD 21
 Qy 1 KPAPWPKVCGGRRLSDQEQSQTETP 17

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:13:48 ; Search time 16.8269 Seconds
(without alignments)
142.913 Million cell updates/sec

Title: US-10-008-524a-4

Perfect score: 142

Sequence: 1.RPIKPSWAPKXRRRLSDQSQTP 25

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched:

283366 seqs, 9619526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database :

PIR 7B.*

1: pir1;*

2: pir2;*

3: pir3;*

4: pir4;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	121	85.2	95	1	W4WLHS		B4 protein - human
2	80	56.3	96	1	W4WL35		E4 protein - human
3	74	52.1	87	1	W4WL51		E4 protein - human
4	70.5	49.6	102	1	W4WL31		E4 protein - human
5	61.5	43.3	1553	2	T03301		rabb effector Prot
6	59	41.5	94	1	W4WL39		E4 protein - human
7	57	40.1	230	2	S72714		Lerchh170 F2-64 pro
8	56	39.4	137	2	T43028		HMG-protein I alph
9	54	38.0	330	2	T05717		Protein extensin
10	53	37.3	88	1	W4WL18		E4 protein - human
11	53	37.3	238	2	T40820		proline-rich Prote
12	53	37.3	275	2	DNA-directed RNA P		DNA-directed RNA P
13	53	37.3	910	2	JC4609		chitin synthase (E
14	52.5	37.0	684	2	H96646		hypothetical Prote
15	52.5	37.0	684	2	T02149		hypothetical Prote
16	52	36.6	459	2	S03116		gene 33 protein, h
17	52	36.6	1223	2	T17345		hypothetical Prote
18	51	35.9	101	2	S30970		gene 25 protein -
19	51	35.9	1039	2	A34269		integrin alpha-2b
20	51	35.9	1060	2	AI0201		beta-galactosidase
21	50.5	35.6	327	2	T45143		hypothetical Prote
22	50	35.2	38	1	HSPY5		histone H5 - Piego
23	50	35.2	105	1	NSCHH4		nonhistone chromos
24	50	35.2	338	2	C75459		probable endonuclease
25	49.5	34.9	529	2	A47082		probable transcripti
26	49.5	34.9	529	2	AI2119		transcription regu
27	49.5	34.9	838	2	T04785		hypothetical Prote
28	49.5	34.9	1487	2	T02850		hypothetical Prote
29	49	34.5	101	2	G72802		gp25 protein - Myc

RESULTS

W4WLHS

B4 protein - human

E4 protein - human

E4 protein - human

rabb effector Prot

E4 protein - human

Lerchh170 F2-64 pro

HMG-protein I alph

Protein extensin

E4 protein - human

proline-rich Prote

DNA-directed RNA P

chitin synthase (E

hypothetical Prote

hypothetical Prote

gene 33 protein, h

hypothetical Prote

gene 25 protein -

integrin alpha-2b

beta-galactosidase

hypothetical Prote

histone H5 - Piego

nonhistone chromos

probable endonuclease

probable transcripti

transcription regu

hypothetical Prote

hypothetical Prote

gp25 protein - Myc

RESULT 1

W4WLHS

B4 protein - human

E4 protein - human papillomavirus type 16

C;Species: human papillomavirus type 16

C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 20-Aug-1999

R;Accession: A2355; T10425

R;Sedorff, K.; Kramer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.

Virology 145, 181-85, 1985

A;Title: Human papillomavirus type 16 genome acts at the lev

A;Reference number: A22355; PMID:2990099

A;Accession: A22355

A;Molecule type: DNA

A;Residues: 1-95 <GEN>

A;Cross-references: EMBL:KO2718; NID:9333031; PIDN:AAA46937.1; PID:9459913

R; Kennedy, I.M.; Haddow, J.K.; Clements, J.B.

J. Virol. 65, 2093-2097, 1991.

A;Title: A negative element in the human papillomavirus type 16 genome acts at the lev

A;Reference number: Z17014; PMID:9116263; PMID:1848319

A;Accession: T10425

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-95 <GEN>

A;Cross-references: EMBL:KO2718; NID:9333031; PIDN:AAA46937.1; PID:9459913

C;Genetics:

A;Gene: B4

C;Superfamily: papillomavirus E4 protein

C;Keywords: early protein

Query Match

Best Local Similarity

Matches 25; Conservative

0; Mismatches 0;

Indels 2; Gaps 2;

Db 29 RPLPKSPNPKKXRL-SDQD-SQTP 25

Db 29 RPLPKSPNPKKXRL-SDQD-SQTP 55

RESULT 2

W4WL35

B4 protein - human papillomavirus type 35

C;Species: human papillomavirus type 35

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 27-Jan-1995

R;Accession: C40824

R;Marich, J.E.; Pontsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubinsky, T.W.

Virology 186, 770-776, 1992

A;Title: The phylogenetic relationship and complete nucleotide sequence of human papill

A;Reference number: A40824; PMID:92124753; PMID:1310198

A;Accession: C40824

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-96 <Myc>

ALIGNMENTS

A;Cross-references: GB:M74117
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 56.3%; Score 80; DB 1; Length 96;
 Best Local Similarity 66.7%; Pred. No. 0.0012; 0; Gaps 0;
 Matches 14; Conservative 2; Mismatches 5; Indels 0;

QY 1 RPIPKSPWAKKHKRRLSDQD 21
 DB 29 RPIPKSPWAKPQKPRRQDITND 49

RESULT 3
 W4WL51
 E4 protein - human papillomavirus type 51.
 C;Species: human papillomavirus type 51
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 27-Jan-1995
 C;Accession: E0415
 R;Lungu, O.; Crum, C.P.; Silverstein, S.J.
 J. Virol. 65, 4216-4225, 1991
 A;Title: Biologic properties and nucleotide sequence analysis of human papillomavirus type 51.
 A;Reference number: A40415; PMID:91303675;
 A;Accession: C40415
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-87 <LNU>
 A;Cross-references: GB:M62877
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 52.1%; Score 74; DB 1; Length 87;
 Best Local Similarity 63.6%; Pred. No. 0.0065; 0; Gaps 0;
 Matches 14; Conservative 1; Mismatches 7; Indels 0;

QY 1 RPIPKSPWAKKHKRRLSDQD 22
 DB 24 RPIPLPAWAKPQPRNSEND 45

RESULT 4
 W4WL31
 E4 protein - human papillomavirus type 31.
 C;Species: human papillomavirus type 31
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
 C;Accession: E32444
 R;Goldsbrough, M.D.; Disilvestre, D.; Temple, G.F.; Lorincz, A.T.
 Virology 171, 306-311, 1989
 A;Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated virus
 A;Reference number: A49398; PMID:89299478;
 A;Accession: E32444
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-102 <SQL>
 A;Cross-references: GB:J04353; NID:9333048; PIDN:AAA46949.1; PID:9459915
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 49.6%; Score 70.5; DB 1; Length 102;
 Best Local Similarity 50.0%; Pred. No. 0.022; 3; Gaps 2;
 Matches 16; Conservative 4; Mismatches 3; Indels 9;

QY 3 IPKPSWAKK---HRRLSQDQ---SQTP 25
 DB 32 IPKPAPAPVYCGGRLLSDEQSOSTETP 63

RESULT 7
 S72714
 Lepb1170_F2_64 protein - Mycobacterium leprae
 C;Species: Mycobacterium leprae
 C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C;Accession: S72714
 R;Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A;Description: Mycobacterium leprae cosmid B1170.
 A;Reference number: S72693
 A;Accession: S72714
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-30 <SMI>
 A;Cross-references: EMBL:U00010; NID:9466780; PIDN:AAA17078.1; PID:9466802
 C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1517

RESULT 5
 T03301
 tab3 effector protein Rim - rat

Query Match 40.1%; Score 57; DB 2; Length 230;
Best Local Similarity 57.9%; Pred. No. 2.9;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
RESULT 10

W4WL18
E4 protein - human papillomavirus type 18
C;Species: human papillomavirus type 18
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
C;Accession: B26251
A;Accession: B26251
A;Molecule type: DNA
R;Cole, S.T.; Danos, O.
J. Mol. Biol. 193, 599-608, 1987
A;Title: Nucleotide sequence and comparative analysis of the human papillomavirus type
A;Reference number: A92937; MUID:8728382; PMID:3039145

RESULT 8

T43028
HMG-protein I alpha chain - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Accession: T43028; T16801
R;Kurz, T.; Schulze, E.
submitted to the EMBL Data Library, April 1998
A;Description: The high mobility group proteins of *Caenorhabditis elegans*.
A;Reference number: Z22282
A;Accession: T43028
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-137 <CHI>
A;Cross-references: EMBL:U40028; NID:gi1055143; PIDN:AAA81116.1; CESP:T05A7
C;Genetics:
A;Gene: hmgi-I-alpha; T05A7.4
A;Introns: 29/1

Query Match 39.4%; Score 56; DB 2; Length 137;
Best Local Similarity 44.0%; Pred. No. 2.4;
Matches 11; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
RESULT 9

QY 1 RPIPKPSWAPKKHRLSDQSTP 25
Db 53 KPKPKPSGAGPRGRPKSKDQVAKSP 77

Query Match 38.0%; Score 54; DB 2; Length 330;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
RESULT 11

T40820
proline-rich protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: T40820
R;Beck, A.; Reinhardt, R.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, October 1998
A;Accession: T40820
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-238 <BEC>
A;Cross-references: EMBL:AL032684; PIDN:CAA21811.1; GSPDB:GN00067; SFDB:SPBP8BT.26
A;Experimental source: strain 972h-1; clone p1 p8B7
C;Genetics:
A;Gene: SPBP8BT.26
A;Map Position: 2
C;Superfamily: proline-rich protein

Query Match 37.3%; Score 53; DB 2; Length 238;
Best Local Similarity 52.4%; Pred. No. 10;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
RESULT 12

QY 2 PIPKPSWAPKKHRLSDQSTP 22
Db 194 PVKKDPPPAPKPKPRLAARTS 214

Query Match 38.0%; Score 54; DB 2; Length 330;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
RESULT 13

T05717
probable extensin - barley (fragment)
C;Species: Hordeum vulgare (barley)
C;Accession: T05717
R;Doan, D.N.P.; Stararo, M.; Olsen, O.A.
submitted to the EMBL Data Library, July 1997
A;Description: Characterization of a nucellar cDNA encoding a probable extensin from developing seed
A;Accession: T05717
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-330 <DOA>
A;Cross-references: EMBL:Z98204; PIDN:CAB10894.1
C;Genetics:
A;Gene: ex1
A;Map position: 2
C;Superfamily: hydroxyproline-rich glycoprotein

Query Match 38.0%; Score 54; DB 2; Length 330;
Best Local Similarity 75.0%; Pred. No. 10;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
RESULT 14

QY 1 RPIPKPSWAPK 12
Db 200 KPVPKPSPAFK 211

A;Accession: A36264
DNA-directed RNA polymerase (EC 2.7.7.6) II 33K chain - human
C;Species: Homo sapiens (man)
C;Accession: A36264
R;Pati, U.K.; Weissman, S.M.
J. Biol. Chem. 265, 8400-8403, 1990
A;Title: The amino acid sequence of the human RNA polymerase II 33-kDa subunit hRPB 33
A;Reference number: A36264; MUID:90256750; PMID:2187864
A;Accession: A36264
A;Molecule type: tRNA
A;Residues: 1-275 <PAT>
A;Cross-references: GB:J05448; NID:g337496; PIDN:AAA3586.1; PMID:937497
C;Superfamily: DNA-directed RNA Polymerase II chain RPB3; ferredoxin 2 [4Fe-4S] homology

C;Keywords: nucleotidyltransferase; transcription
 Query Match 37.3%; Score 53; DB 2; Length 275;
 Best Local Similarity 36.4%; Pred. No. 12;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy 4 PKPSPWPKKHLRLLSDQDSQTP 25
 Db 1.98 PKPEENPKSEYSELDEESQAP 219

RESULT 15
 T02149
 hypothetical protein F8K4_24 - Arabidopsis thaliana (fragment)
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 05-Mar-1999 #text_change 22-Oct-1999
 C;Accession: T02149
 R;Vysotskaya, V.S.; Schwartz, J.R.; Toriumi, M.; Kwan, A.; Yu, G.; Oji, O.; Liu, S.; Liu, R.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Eckner, J.R.; Federspiel, N.; Weiss, N.; Szebenyi, A.; Yarden, O.
 Gene 168, 99-102, 1996
 A;Title: The chsa gene, encoding a class-I chitin synthase from Ampelomyces quisqualis.
 B;Reference number: JC4609; MUID:96186963; PMID:8626074
 A;Accession: JC4609
 A;Molecule type: DNA
 B;Residues: 1-910 <WRI>
 A;Cross-references: EMBL:X86802; NID:gl1406268; PIDN:CAA60497.1; PID:91486269
 C;Comment: This enzyme belongs to the class-I chitin synthases, and is a member of a multi-gene family.
 A;Gene: chsa
 A;Introns: 324/3
 C;Function:
 A;Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosamine transferase
 C;Superfamily: chitin synthase chSA
 C;Keywords: glycosyltransferase, hexosyltransferase

Query Match 37.3%; Score 53; DB 2; Length 910;
 Best Local Similarity 60.0%; Pred. No. 39;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PKPSPWPKKHLRLLSDQDSQTP 25
 Db 7 PSRSPWPAFTQRRLA 21

RESULT 14
 H96646
 hypothetical protein F8K4_24 [Imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #text_change 31-Mar-2001
 C;Accession: H96646
 P;Theologis, A.; Eckner, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chang, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansen, N.F.; Hughes, B.; Huijzer, L.; Nature 408, 916-920, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luero, J.S.; Maiti, R.; Marzali, P.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Talion, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: H96646
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-684 <SPO>
 A;Cross-references: GB:AE005173; NID:gl3367537; PIDN:AAC28522.1; GSPDB:GN00141
 C;Genetics:
 A;Gene: FB4_24
 A;Map position: 1

Query Match 37.0%; Score 52.5; DB 2; Length 684;
 Best Local Similarity 35.7%; Pred. No. 34;
 Matches 10; Conservative 5; Mismatches 8; Indels 5; Gaps 1;

Scoring table:	BLOSUM62	ALIGMENTS				
Gapop:	10.0 , Gappext 0.5					
Searched:	141681 seqs, 52070155 residues					
Total number of hits satisfying chosen parameters:	141681					
Minimum DB seq length:	0					
Maximum DB seq length:	2000000000					
Post-processing:	Minimum Match 0%					
Database :	SwissProt_42:*					
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES						
Result No.	Score	Query	Description			
		Match	Length	DB ID		
- - -	- - -	- - -	- - -	- - -	- - -	- - -
1	121	85.2	95	1 VE4 HPV16	P06222 human papill	RT
2	80	56.3	95	1 VE4 HPV35	P27224 human papill	RT
3	74	52.1	87	1 VE4 HPV51	P26548 human papill	CC
4	70	49.6	102	1 VE4 HPV31	P1784 human papill	CC
5	61.5	43.3	1615	1 RIM1 RAT	Q9jxr4 rat	CC
6	61.5	43.3	1692	1 RIM1_HUMAN	Q86rs3 homo sapien	CC
7	59	41.5	94	1 VE4 HPV39	P24311 human papill	CC
8	54	38.0	670	1 ZN16_HUMAN	P17620 human papill	CC
9	53	37.3	88	1 VE4 HPV18	P06791 human papill	CC
10	53	37.3	275	1 RPB3_HUMAN	P19387 homo sapien	CC
11	53	37.3	755	1 MTS1_HUMAN	O43112 homo sapien	CC
12	53	37.3	910	1 NSAA_AMPOU	Q12654 amelonyces	DR
13	52.5	37.0	1530	1 RIM2_MOUSE	Q9ez77 mus musculus	DR
14	52	36.6	459	1 MIG6_RAT	P05332 rattus norvegicus	DR
15	52	36.6	461	1 MIG6_MOUSE	Q99127 mus musculus	DR
16	52	36.6	1508	1 GEM5_HUMAN	Q8tq96 homo sapien	DR
17	51	35.9	101	1 VG25_BPM15	Q05332 mycobacterium	DR
18	51	35.9	759	1 MTS1_MOUSE	Q8r184 mus musculus	DR
19	51	35.9	1039	1 ITAB_HUMAN	P0814 homo sapien	DR
20	50.5	35.6	1142	1 ENAM_HUMAN	Q9rrml homo sapien	DR
21	50	35.2	38	1 H5_COLLI	P02660 columba livia	DR
22	50	35.2	104	1 HG15_CHICK	P1202 gallus gallus	DR
23	49.5	34.9	361	1 IHA_TRIVU	Q77555 trichosurus	DR
24	49.5	34.9	1188	1 RIM2_HUMAN	Q9uq26 homo sapien	DR
25	49.5	34.9	1555	1 RIM2 RAT	Q9jjs1 rattus norvegicus	DR
26	49	34.5	101	1 VG25_BPM12	Q64219 mycobacterium	DR
27	49	34.5	135	1 INI5_HUMAN	Q9yq96 homo sapien	DR
28	49	34.5	325	1 CTH1_YEAST	P47766 saccharomyces	DR
29	49	34.5	393	1 TRMB_HELPPY	P25443 helicobacte	DR
30	49	34.5	400	1 TRMB_HELPY	Q9z196 helicobacte	DR
31	49	34.5	893	1 YM92_GAEEL	P34531 caenorhabdi	DR
32	49	34.5	956	1 YEF3_YEAST	P32618 saccharomyces	DR
33	49	34.5	1207	1 DML1_ARATH	Q9sj96 arabidopsis	DR
					RESULT 2	
					VE4 HPV15	
					ID: P07455	
					STANDARD;	
					PRT;	
					96 AA.	
					AC: P27224;	
					DT: 01-AUG-1992 (Rel. 23, Created)	

DT	01-AUG-1992 (Rel. 23, Last sequence update)	Query Match Score 74; DB 1; Length 87;
DE	Probable E4 protein.	Best Local Similarity 63.6%; Pred. No. 0.006; Gaps 0;
CDS	Human papillomavirus type 35.	Matches 14; Conservative 1; Mismatches 7; Indels 0;
CC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;	
CC	Papillomavirus.	
CX	NCBI_TaxID=10587;	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=92124753; PubMed=1310198;	
RA	Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;	
RT	"The phylogenetic relationship and complete nucleotide sequence of human papillomavirus type 35.",	
RT	Virology 186:770-776(1992).	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	EMBL; M74117; AAA44695.2; -.	
DR	PIR; C40824; W4W135.	
DR	InterPro; IPR003861; Papilloma_E4.	
DR	Pfam; PF02711; Pap_E4; 1.	
FW	Early protein.	
SQ	SEQUENCE 96 AA; 10597 MW; AB4524418CD26F7C CRC64;	
Query Match Score 80; DB 1; Length 96;		
Best Local Similarity 66.7%; Pred. No. 0.0011; Gaps 0;		
Matches 14; Conservative 2; Mismatches 5; Indels 0; Gaps 0;		
CC	Probable E4 protein.	
CC	Human papillomavirus type 51.	
CC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;	
CC	Papillomavirus.	
CX	NCBI_TaxID=10595;	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=9103675; PubMed=1649326;	
RA	Lungu O., Crum C.P., Silverstein S.J.;	
RT	"Biological properties and nucleotide sequence analysis of human papillomavirus type 51";	
RT	J. Virol. 65:4216-4225(1991).	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	EMBL; M62877; -; NOT_ANNOTATED_CDS.	
DR	PIR; C40415; W4W151.	
DR	InterPro; IPR003861; Papilloma_E4.	
DR	Pfam; PF02711; Pap_E4; 1.	
FW	Early protein.	
SQ	SEQUENCE 87 AA; 9941 MW; 5F3DC38E86BF3990 CRC64;	
Query Match Score 74; DB 1; Length 87;		
Best Local Similarity 63.6%; Pred. No. 0.006; Gaps 0;		
Matches 14; Conservative 1; Mismatches 7; Indels 0; Gaps 0;		
CC	Probable E4 protein.	
CC	Human papillomavirus, no RNA stage; Papillomaviridae;	
CC	Papillomavirus.	
CX	NCBI_TaxID=10587;	
RP	SEQUENCE FROM N.A.	
RX	(ISOFORM 2), AND INTERACTION WITH RAB3A AND RAB3C.	
RC	TISSUE=Brain;	
RX	MEDLINE=97394473; PubMed=9252191;	
RA	Wang Y., Okamoto M., Schmitz F., Hofmann K., Suedhof T.C.;	

"Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion.";

RT Nature 388:593-598(1997).

RN [2]

SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.

RC TISSUE=Brain;

RC MEDLINE=;

RA WANG Y., Sugita S., Sudhof T.C., "The RIM/NIM Family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins.", J. Biol. Chem. 275:20033-20044 (2000)

-!- FUNCTION: Rab effector involved in exocytosis. May act as scaffold protein.

CC -!- SUBUNIT: Binds SNAP25, SYT1 and CACNA1B. Interaction with SYT1 is enhanced by calcium ions. Interaction with SNAP25 is weaker in the presence of calcium ions. Binds RAB3A, RAB3B and RAB3D that have been activated by GTP-binding. Binds UNC13 (By similarity).

CC -!- SUBCELLULAR LOCATION: peripheral membrane protein associated with plasma membranes from synaptic junctions. Not detected in synaptic vesicles. Detected in presynaptic nerve terminals close to the active zone. Detected in synaptic ribbons of ribbon synapses of retinal photoreceptor cells.

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=2; IsoId=Q9JTR4;1; Sequence=Displayed;

CC -!- TISSUE SPECIFICITY: Highly expressed in hippocampus, brain cortex, cerebellum and olfactory bulb. Detected at lower levels in midbrain, hindbrain and spinal cord. Detected retina and in spinal cord motor neurons.

CC -!- SIMILARITY: Contains 2 C2 domains.

CC -!- SIMILARITY: Contains 1 FYVE-type zinc finger.

CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.

CC -!- SIMILARITY: Contains 1 Rab-binding (RabBD) domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AF007836; AB66703.1; -.

DR PIR; T03301; T03301.

DR HSSP; P21707; 1RSY.

DR InterPro; IPR000893; C2_CalB.

DR InterPro; IPR0101478; PDZ.

DR InterPro; IPR003315; RPH3A_effector.

DR InterPro; IPR000306; Znf_FYVE.

DR Pfam; PF00168; C2_2.

DR Pfam; PF00595; PDZ_1.

DR Pfam; PF02318; RPH3A_effector; 1.

DR SMART; SM00239; C2_2.

DR PROSITE; PS50004; C2_DOMAIN_2; 2.

DR PROSITE; PS50106; PDZ; 1.

DR PROSITE; PS50316; RABBD_1.

DR PROSITE; PS50178; ZF_FYVE; 1.

KW Metal-binding; Zinc_Repeat; Zinc-finger; Alternative splicing.

FT DOMAIN 22 205 RAB-BINDING.

FT DOMAIN 42 2 193 FYVE-TYPE.

FT DOMAIN 44 455 ALA-RICH.

FT DOMAIN 44 483 PRO-RICH.

FT DOMAIN 619 705 PDZ.

FT DOMAIN 758 864 C2 DOMAIN 1.

FT DOMAIN 1259 1301 SER-RICH.

FT DOMAIN 1461 1563 C2 DOMAIN 2.

FT VARRSPHIC 1107 1168 Missing (in isoform 2).

FT FT /FTId=VSP_008172;

SQ SEQUENCE 1615 AA; 17654 MW; 80E76F74BF5FB7E CRCC44;

Query Match 43.3%; Score 61.5; DB 1; Length 1615;

Best Local Similarity 33.3%; Pred. No. 3-8;

Matches 11; Conservative 7; Missmatches 2; Indels 13; Gaps 1;

Qy 2 PIPKSPWPKH-----RLSDQD 21

Db 890 PLPQFSPFMFRHIGESSSSKKLQERSQRISDSD 922

RESULT 6

RIM1_HUMAN STANDARD PRT_1692_AA

ID RIM1_HUMAN STANDARD PRT_1692_AA

AC Q86T5; Q15048; Q8TDY9; Q8TDZ5; Q9HBA1; Q9HBA2; Q9HBA3; Q9HBA4;

AC Q9HBA6;

DT 10-OCT-2003 (Rel. 42; Created)

DT 10-OCT-2003 (Rel. 42; Last sequence update)

DT 15-APR-2004 (Rel. 43; Last annotation update)

DE Regulating synaptic membrane exocytosis protein 1 (Rab3-interacting molecule 1) (RIM 1).

DE RIM1 OR RIM1 OR KIA0340.

GN OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Buteraria; Primates; Catarrhini; Homidae; Homo.

OC NCBI_TaxID=9606;

OX RN [1]

RP SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPlicing, AND VARIANT CORD7 HIS-820.

RX MEDLINE=12545706; PubMed=12659814;

RA RA Johnson S., Halford S., Morris A.G., Patel R.J., Wilkie S.E., Harcastle A.J., Moore A.T., Zhang K., Hunt D.M., RT "Genomic organisation and alternative splicing of human RIM1," a gene implicated in autosomal dominant cone-rod dystrophy (CORD7)."; RT Genomics 81:304-314 (2003).

RL RN [2]

RP SEQUENCE FROM N.A. (ISOFORMS 2 AND 3).

RC TISSUE=Neuroblastoma;

RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Nomura N., Ohara O., Kato T., Nakagawa A.; RT "Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

RA RL DNA Res. 4:141-15(1997).

RN RN [3]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC TISSUE=Brain;

RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Nomura N., Ohara O., Kato T., Nakagawa A.; RT "Identification of the alternative form of human RIM1."; RT Submitted (Nov-2000) to the EMBL/GenBank/DBJ databases.

RA RL DNA Res. 4:141-15(1997).

RN RN [4]

RP SEQUENCE OF 776-1692 FROM N.A. (ISOFORMS 3; 4; 5; 6; 7 AND 8).

RC TISSUE=Brain;

RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N., Miyajima N., Tanaka A., Nomura N., Ohara O., Kato T., Nakagawa A.; RT "Prediction of the coding sequences of unidentified human genes. VII. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

RA RL DNA Res. 4:141-15(1997).

RP MUTAGENESIS OF 796-LYS-LYS-197 AND 1591-ARG-ARG-1592, AND INTERACTION WITH SNAP25; SYT1 AND CACNA1B.

RC TISSUE=Brain;

RA Coppola T., Magnin-Luethi S., Perret-Menoud V., Gattesco S., Schiavo G., Regazzi R.; RT "Direct interaction of the Rab3 effector RIM with Ca2+ channels, SNAP-25, and synaptotagmin."; RT 276:32756-32762(2001).

RL J. Biol. Chem. 276:32756-32762(2001)

CC -!- FUNCTION: Rab effector involved in exocytosis. May act as scaffold protein.

CC CC GTP-binding. Binds UNC13 (By similarity). Binds SNAP5, SYT1 and CACNA1B. Interaction with SYT1 is enhanced by calcium ions.

CC CC Interaction with SNAP25 is weaker in the presence of calcium ions.

CC CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=8;

Comment=Additional isoforms seem to exist;

Name=_1; IsoId=Q86TR5-1; Sequence=Displayed;

Name=_2; Synonyms=RIM short form; Sequence=VSP_008160, VSP_008165;

Name=_3; Synonyms=RIM long form; Rab3 interacting protein variant 2;

IsoId=Q86TR5-3; Sequence=VSP_008160, VSP_008164, VSP_008167,

VSP_008171;

Name=_4; Synonyms=Rab3 interacting protein variant 1;

IsoId=Q86TR5-4; Sequence=VSP_008166, VSP_008167;

Name=_5; Synonyms=Rab3 interacting protein variant 3;

IsoId=Q86TR5-5; Sequence=VSP_008161, VSP_008163, VSP_008170;

Name=_6; Synonyms=Rab3 interacting protein variant 4;

IsoId=Q86TR5-6; Sequence=VSP_008164, VSP_008168, VSP_008169;

Name=_7; Synonyms=Rab3 interacting protein variant 5;

IsoId=Q86TR5-7; Sequence=VSP_008161, VSP_008164, VSP_008167,

VSP_008169;

Name=_8; Synonyms=Rab3 interacting protein variant 6;

IsoId=Q86TR5-8; Sequence=VSP_008161, VSP_008162, VSP_008169;

-- TISSUE SPECIFICITY: Detected in brain and retina.

-!- DISEASE: Defects in RIMS1 are a cause of autosomal dominant cone-rod dystrophy (CORD7) [MIM:60369]. CORD7 is characterized by early loss of visual acuity and color vision, followed by night blindness and peripheral visual field loss. The onset of reduced color vision and visual acuity varies between the ages of 20 and 40 years.

-!- SIMILARITY: Contains 2 C2 domains.

-!- SIMILARITY: Contains 1 FYVE-type zinc finger.

-!- SIMILARITY: Contains 1 PDZ/DHR domain.

-!- SIMILARITY: Contains 1 Rab-binding (RabBD) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AY190519; AA038848.1; .

DR EMBL; AB045726; BABBT121.1; .

DR EMBL; AB051866; BABBT242.1; .

DR EMBL; AB02338; BAA20798.1; ALT_INIT..

DR EMBL; AF263305; AAG23162.1; ALT_INIT..

DR EMBL; AF263306; AAG23163.1; .

DR EMBL; AF263307; AAG23164.1; .

DR EMBL; AF263308; AAG23165.1; .

DR EMBL; AF263309; AAG23166.1; .

DR EMBL; AF263310; AAG23167.1; .

DR HSSP; P21707; IRSY.

DR GeneW HGNC:17222; RIMS1.

DR MIM; 606629; .

DR MIM; 603649; .

DR InterPro; IPR000008; C2.

DR InterPro; IPR000973; C2_CalB.

DR InterPro; IPR001478; PDZ.

DR SMART; SM00239; C2_2.

DR PFam; SM00228; PDZ_1.

DR PFam; PF00168; C2_2.

DR PFam; PF00595; PDZ_1.

DR PFam; PF02318; RPH3A_effector_1.

DR SMART; SM00239; C2_2.

DR PROSITE; PS50004; C2_DOMAIN_2; .

DR PROSITE; PS50106; PDZ_1.

DR PROSITE; PS50116; RAB3D_1.

DR PROSITE; PS50178; ZF_FYVE_1.

KW Metal-binding; Zinc; Repeat; Zinc-finger; Alternative splicing;

KW Vision; Disease mutation.

FT DOMAIN 22 182 RAB-BINDING.

FT ZN_FINGER 110 170 FYVE-TYPE.

FT DOMAIN 691 605 PDB.

FT DOMAIN 744 850 C2 DOMAIN 1.

FT DOMAIN 1538 1640 C2 DOMAIN 2.

FT DOMAIN 1336 1402 SER-RICH.

FT VARSPLIC 484 494 Missing (in isoform 2 and isoform 3).

FT VARSPLIC 924 924 Missing (in isoform 5, isoform 7 and isoform 8).

FT VARSPLIC 1018 1245 Missing (in isoform 8).

FT VARSPLIC 1038 1244 /FTID=VSP_008162.

FT VARSPLIC 1039 1102 Missing (in isoform 5).

FT VARSPLIC 1040 1692 Missing (in isoform 2).

FT VARSPLIC 1065 1102 /FTID=VSP_008163.

FT VARSPLIC 1133 1245 Missing (in isoform 3, isoform 6 and isoform 7).

FT VARSPLIC 1161 1245 /FTID=VSP_008164.

FT VARSPLIC 1284 1455 Missing (in isoform 6, isoform 7 and isoform 8).

FT VARSPLIC 1377 1385 /FTID=VSP_008165.

FT VARSPLIC 1540 1573 Missing (in isoform 3).

FT VARIANT 820 820 /FTID=VSP_008171.

FT MUTAGEN 796 797 R->H (in R168D).

FT MUTAGEN 1591 1592 RR->A; ABOLISHES INTERACTION WITH SYT1 AND CACNA1B.

SEQUENCE 1692 AA; 189071 MW; DA96642DCB32C15E CRC64;

Query Match 43.3%; Score 61.5; DB 1; Length 1692;

Best Local Similarity 33.3%; Pred. No. 4; Matches 11; Conservative 7; Wimatches 2; Indels 13; Gaps 1;

Qy 2 PIPKPSWPWAKKH-----RRLSQD 21

Db 876 PLPQPSPFMPERPHIGESSSSKRQLQRSGRISDSD 908

RESULT 7

VE4_HPV39 STANDARD; PRT; 94 AA.

ID VE4_HPV39

AC P24831;

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 01-MAR-1992 (Rel. 21, Last annotation update)

DE Probable E4 protein.

OS Human Papillomavirus type 39.

OC Papillomavirus, no RNA stage; Papillomaviridae;

OC Viruses; dsDNA viruses;

OX NCBI_TaxID:10588;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91135017; PubMed=1847266;

RX "Volpers C., Streck R.E.;

RX Virology 181:419-423 (1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

C entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
C or send an email to license@isb-sib.ch)

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sis.ch/announce/> or send an email to license@isb-sis.ch).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - The European Bioinformatics Institute. There are no restrictions on its use.

CC modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@isb-sib.ch).

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustini T.B., Tostiyuki S., Carninci P., Prange C., RA Raha S.S., Lognelli N.A., Peters G.J., Abramson R.D., Mullahy S.J., RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley S.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W., RA Villalobos D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., RA Fahey J., Heitman E., Ketteman M., Madan A., Rodrigues S., Sanchez A., RA Whiting M., Madan A.C., Young B., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmitz J., Myers R.M., RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalius D.E., RA Schncher A., Schein J.E., Jones S.J.M., Marras M.A.; RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). CC -1- FUNCTION: DNA-dependent RNA polymerase catalyzes the transcription of DNA into RNA using the four ribonucleoside triphosphates as substrates. CC -1- CATALYTIC ACTIVITY: Nucleoside triphosphate = N diphosphate + [RNA] (N). CC -1- SUBUNIT: RNA POLYMERASE II CONSISTS OF 12 DIFFERENT SUBUNITS. CC -1- POLR2J AND POLR2C SUBUNITS INTERACT WITH EACH OTHER. CC -1- SUBCELLULAR LOCATION: Nuclear. CC -1- MISCELLANEOUS: Three distinct zinc-containing RNA polymerases are found in eukaryotic nuclei: polymerase I for the ribosomal RNA precursor, polymerase II for the mRNA precursor, and polymerase III for 5S and tRNA genes.

CC -1- SIMILARITY: Belongs to the archaeabacteria RPOD / eukaryotic RPB3 RNA polymerase subunit family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; J05448; AAA36386.1; -. DR EMBL; AF008443; AAC14355.1; -. DR EMBL; AJ224143; CAA11842.1; -. DR EMBL; AJ224144; CAA11843.1; -. DR EMBL; AC004382; AAC24309.1; -. DR EMBL; BC000409; AAH00409.1; -. DR EMBL; BC003159; AAH03159.1; -. DR EMBL; BC028157; AAH28157.1; -. DR PIR; A36264; A36264. DR Genew; HGNC; 9189; POLR2C. DR MIM; 180653; -. DR GO; GO:0006366; P:transcription from Pol II promoter; TAS. DR InterPro; IPR001700; RNA_pola_bac_org. DR InterPro; IPR001514; RNA_PoLD. DR Pfam; PF0100; RNA_DolA_bac; 1. DR ProDom; PD002883; RNA_PoLD; 1. DR SMART; SM00662; RPOD; 1. DR PROSITE; PS00446; RNA_PoLD_30KD; 1. DR Transferase; DNA-directed RNA polymerase; KW Nuclear Protein. DR DOMAIN 88 97 CYS-RICH. DR CONFFLICT 194 194 H -> T (IN REF. 1 AND 3). DR SEQUENCE 275 AA; 31441 MW; EF663BEE964A4B CRC64; DR Query Match 37.3%; Score 53; DB 1; Length 275; Best Local Similarity 36.4%; Pred. No. 8.3; Mismatches 4; Indels 0; Gaps 0;

QY 4 PKPSPWPKHRLSPQDSQP 25

Db 198 PKPPEPKSEYSELDEDESQF 219

MTS1_HUMAN	STANDARD;	PRT;	755 AA.	
AC	Q43722; Q8TC22; Q96TX2;			
DT	16-OCT-2001 (Rel. 40; Created)			
DT	10-OCT-2003 (Rel. 42; Last sequence update)			
DT	10-OCT-2003 (Rel. 42; Last annotation update)			
DE	Metastasis suppressor protein 1 (Missing in metastasis protein) (Metastasis suppressor protein 1 (Missing in metastasis protein))			
DE	MTSS1 OR MIM OR KIAA0429.			
OS	Homo sapiens (Human);			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX				
RN	[1] RP SEQUENCE FROM N.A. (ISOFORM 1).			
RA	Lee Y.-G., Macoska J.A., Schwab E.D., Korenchuk S., Pienta K.J.;			
RT	"Identification of potential metastasis suppressor gene (YCH-1) in bladder cancer."			
RT	"NEDO human cDNA sequencing project."			
RL	Submitted (AUG-1998) to the EMBL/GenBank/DBJ/GenBank/DBJ databases.			
RN	[2] RP SEQUENCE FROM N.A. (ISOFORM 1).			
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshihikawa Y., Matsunawa H., Ishii S., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahashi K., Masutoh Y., Sasaki N.,			
RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RT	"Prediction of the coding sequences of unidentified human genes. VIII. RT 78 new cDNA clones from brain which code for large proteins in vitro."			
RL	Submitted (SEP-2000) to the EMBL/GenBank/DBJ/GenBank/DBJ databases.			
RN	[3] RP SEQUENCE FROM N.A. (ISOFORM 3).			
RC	TISSUE=Brain;			
RX	MEDLINE=90116655; PubMed=9455477;			
RA	Ishikawa N., Tanaka A., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;			
RA	"Prediction of the coding sequences of unidentified human genes. VIII. RT 78 new cDNA clones from brain which code for large proteins in vitro."			
RL	DNA Res. 4:307-313 (1997).			
RN	[4] RP SEQUENCE FROM N.A. (ISOFORM 3), AND TISSUE SPECIFICITY.			
RX	MEDLINE=2207764; PubMed=12083544;			
RA	Lee Y.-G., Macoska J.A., Korenchuk S., Pienta K.J.;			
RT	"MIM - a potential metastasis suppressor gene in bladder cancer.";			
RL	Neoplasia 4:291-294 (2002).			
RN	[5] RP SEQUENCE FROM N.A. (ISOFORM 2).			
RC	TISSUE=Testis;			
RX	MEDLINE=22388237; PubMed=12477932;			
RA	Straubhaar R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shearmen C.M., Schuler G.D., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Rubin G.M., Hong L., Diachenko L., Marusina K., Farmer A.A., Buetow K.H., Schaefer C.F., Blatner M.J., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Logueillo N.A., Peters G.J., Abramson R.D., Mullally S.J., Bosak S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay J.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Blakesley R.W., Touchman J.W., Green B.D., Bouffard G.G., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E., Schnurch A., Schein J.E., Jones S.J.M., Marra M.A., and mouse cDNA sequences";			
RT	"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
[6] RP BINDING OF WH2 DOMAIN TO ACTIN, AND SUBUNIT.				
RX	MEDLINE=22562833; PubMed=12570871;			
RA	Woodings J.A., Sharp S.J., Machevey L.M.;			
RT	"MIM-B, a putative metastasis suppressor protein, binds to actin and to protein tyrosine phosphatase delta."			
RL	Biochem. J. 371:463-471 (2002).			
CC	-!- FUNCTION: May be related to cancer progression or tumor metastasis in a variety of organ sites, most likely through an interaction with the actin cytoskeleton.			
CC	-!- SUBUNIT: Binds to actin. Binds to the cytoplasmic domain of receptor protein tyrosine phosphatase delta.			
CC	-!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=3;			
CC	Name=1;			
CC	Isoid=043312-1; Sequence=Displayed;			
CC	Name=2;			
CC	Isoid=043312-2; Sequence=VSP 007420; VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	Name=3;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	-!- SIMILARITY: Contains 1 WH2 domain.			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			
CC	Isoid=043312-2; Sequence=VSP 007421;			
CC	Note=No experimental confirmation available;			
CC	NAME=;			
CC	Isoid=043312-3; Sequence=VSP 007419;			
CC	-!- TISSUE SPECIFICITY: Expressed in many tissues, including spleen, thymus, prostate, testis, uterus, colon, and peripheral blood.			
CC	-!- DOMAIN: The WH2 motif at the C-terminus binds to actin monomers.			
CC	NAME=;			
CC	Isoid=043312-1; Sequence=VSP 007420; VSP 007421;			
CC	Name=;			

CC Pleosporales; Leptosphaeraceae; mitosporic Leptosphaeraceae;

OC Ampeleomyces.

OX NCBI_TaxID:50730;

BN SEQUENCE FROM N.A.

RC STRAIN=AQ1;

RX MEDLINE=961886963; PubMed=8626074;

RA Weiss N., Sztejneberg A., Yarden O.;

RT "the chsa gene, encoding a class-I chitin synthase from Ampelomyces quisqualis";

RX Gene 168:99-102(1996).

CC -!- FUNCTION: Plays a major role in cell wall biogenesis.

CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4) - (N-acetyl-

CC beta-D-glucosaminyl) } (N) = UDP + {(1,4) - (N-acetyl-beta-D-

CC glucosaminyl}) (N+1).

CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.

CC -!- SIMILARITY: Belongs to the chitin synthase family. Subfamily class

CC I.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC DR EMBL; X86802; CAA60497.1; -.

DR PIR; JCA4609; JCA4609

DR InterPro; IPR004834; Chitin synth.

DR InterPro; IPR001173; Glyco_trans_2.

DR Pfam; PF01644; Chitin synth_1.

DR ProDom; PD002990; Chitin synth_1.

DR Rfam; RF000001; Multicarbohydrate binding domain.

DR Transmembrane domain.

RX MEDLINE=20512528; PubMed=11056535;

RA Ozaki N., Shibasaki T., Kashina Y., Miki T., Takahashi K., Ueno H.,

RA Sunaga Y., Yano H., Matsuda Y., Iwanaga T., Takai Y., Seino S.;

RP SEQUENCE FROM N.A.

RC STRAIN=AQ1;

RX MEDLINE=961886963; PubMed=8626074;

RA Weiss N., Sztejneberg A., Yarden O.;

RT "the chsa gene, encoding a class-I chitin synthase from Ampelomyces quisqualis";

RX Gene 168:99-102(1996).

CC -!- FUNCTION: Plays a major role in cell wall biogenesis.

CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4) - (N-acetyl-

CC beta-D-glucosaminyl) } (N) = UDP + {(1,4) - (N-acetyl-beta-D-

CC glucosaminyl}) (N+1).

CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.

CC -!- SIMILARITY: Belongs to the chitin synthase family. Subfamily class

CC I.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC DR EMBL; X86802; CAA60497.1; -.

DR PIR; JCA4609; JCA4609

DR InterPro; IPR004834; Chitin synth.

DR InterPro; IPR001173; Glyco_trans_2.

DR Pfam; PF01644; Chitin synth_1.

DR ProDom; PD002990; Chitin synth_1.

DR Rfam; RF000001; Multicarbohydrate binding domain.

DR Transmembrane domain.

RX MEDLINE=20512528; PubMed=11056535;

RA Ozaki N., Shibasaki T., Kashina Y., Miki T., Takahashi K., Ueno H.,

RA Sunaga Y., Yano H., Matsuda Y., Iwanaga T., Takai Y., Seino S.;

RP SEQUENCE FROM N.A.

RC STRAIN=AQ1;

RX MEDLINE=961886963; PubMed=8626074;

RA Weiss N., Sztejneberg A., Yarden O.;

RT "the chsa gene, encoding a class-I chitin synthase from Ampelomyces quisqualis";

RX Gene 168:99-102(1996).

CC -!- FUNCTION: Plays a major role in cell wall biogenesis.

CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4) - (N-acetyl-

CC beta-D-glucosaminyl) } (N) = UDP + {(1,4) - (N-acetyl-beta-D-

CC glucosaminyl}) (N+1).

CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.

CC -!- SIMILARITY: Belongs to the chitin synthase family. Subfamily class

CC I.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC DR EMBL; X86802; CAA60497.1; -.

DR PIR; JCA4609; JCA4609

DR InterPro; IPR004834; Chitin synth.

DR InterPro; IPR001173; Glyco_trans_2.

DR Pfam; PF01644; Chitin synth_1.

DR ProDom; PD002990; Chitin synth_1.

DR Rfam; RF000001; Multicarbohydrate binding domain.

DR Transmembrane domain.

DR

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC	"Rat gene 33; analysis of its structure, messenger RNA and basal promoter activity.";
CC	Nucleic Acids Res. 17:6651-6667(1989).
DR	EMBL; AB021131; BAB18975_1; -.
DR	EMBL; AK032619; BAC27953_1; -.
DR	EMBL; AK084172; BAC38794_1; -.
DR	HSSP; P21107; BYN.
DR	MGD; MGI:2152972; Rims2.
DR	GO:0001515; F:protein binding; IP1.
DR	GO:0017156; P:calcium ion dependent exocytosis; IDA.
DR	GO:0019933; P:cAMP-mediated signaling; IDA.
DR	GO:0031073; P:insulin secretion; IDA.
DR	InterPro; IPR000008; C2.
DR	InterPro; IPR008973; C2_CaB.
DR	InterPro; IPR001378; PDZ.
DR	Pfam; PF00013; Znf_FYVE.
DR	Pfam; PF00168; C2; 2.
DR	Pfam; PF00595; PDZ; 1.
DR	SMART; SM00239; C2; 2.
DR	SMART; SM00228; PDZ; 1.
DR	PROSITE; DS50004; C2_DOMAIN_2; 2.
DR	PROSITE; PS50106; PDZ; 1.
DR	PROSITE; PS50916; RABBD; 1.
DR	PROSITE; PS50178; ZF_FYVE; 1.
KW	Metal-binding; Zinc_Repeat; Zinc-finger; Alternative splicing.
FT	DOMAIN 26 194 RAB-BINDING.
FT	ZN_FING 126 182 FIVE-TYPE.
FT	DOMAIN 677 763 PDZ.
FT	DOMAIN 816 922 C2 DOMAIN 1.
FT	DOMAIN 1376 1478 C2 DOMAIN 2.
FT	DOMAIN 1169 1202 SER-RICH.
FT	VARSPLIC 1 1245 Missing (in isoform 2).
FT	/FTID=VSP /FTID=VSP 008181.
FT	Missing (in isoform 3).
FT	/FTID=VSP /FTID=VSP 008182.
FT	PGDLSVPAVGRAGLRTQDTIKNGSGVKHQIASDMPD->MOPETLRQCVNSVLSPHGVFSSPPNTLQNELFGQTINNA (in isoform 3).
FT	/FTID=VSP 008183.
FT	NDGSQDTAVGALGTSKGRKRSIGKMAVIVGLRSKRSAA
FT	SOLSQQT -> MGROGIGGTGAAGRMORSGRSSLSSASPEA
FT	LAGYEFICMNSLLEDEG (in isoform 2).
FT	/FTID=VSP 008184.
SQ	SEQUENCE 1530 AA; 172863 MW; 58CF11BF7152357D CRC64;
Query Match	Score 52.5%; DB 1; Length 1530;
Best Local Similarity	37.0%; Last sequence update
Matches	27.3%; Pred. No. 50;
Qy	Conservative 8; Mismatches 3; Indels 13; Gaps 1;
DB	2 PIPKPSPWAKKH-----RLSPQD 21 : : : : : : 948 PLPRPSPLPQLHGESPTRLQRSKRISDE 980
RESULT 14	
MG6_RAT	MIG6_RAT STANDARD; PRT; 459 AA.
ID	MIG6_RAT STANDARD; PRT; 459 AA.
AC	P05452;
DT	01-NOV-1988 (Rel. 09, Created)
DT	01-NOV-1988 (Rel. 09, Last sequence update)
DT	28-FEB-2003 (Rel. 41, Last annotation update)
DB	Mitogen-inducible gene 6 protein homolog (Mig-6) (Gene 33 polypeptide).
DE	MIG6 OR 33.
GN	Rattus norvegicus (Rat).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buterria; Rodentia; Sciurognathi; Muridae; Rattus.
OC	NCBI_TaxID=10116;
OX	[1]
RN	SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.
RX	STRAIN=FVB/N; TISSUE=Mammary gland;
ID	MIG6_MOUSE STANDARD; PRT; 461 AA.
AC	Q99z77;
DT	28-FEB-2003 (Rel. 41, Created)
DT	28-FEB-2003 (Rel. 41, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DB	Mitogen-inducible gene 6 protein homolog (Mig-6).
GN	MIG6.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buterria; Rodentia; Sciurognathi; Muridae; Rattus.
OX	NCBI_TaxID=10990;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FVB/N; TISSUE=Mammary gland;
ID	MIG6_MOUSE STANDARD; PRT; 461 AA.
RA	Strandsberg R.L., Peingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shamen C.M., Bhat N.K., Altenschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Gedin I.B., Toshimuki S., Carninci P., Prange C., Raha S.S., Loqueland N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Liu X., Gibbs R.A., Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A.C., Young A.C., Shevchenko Y., Boutin G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krywinski M.I.I., Skalska U., Smailus D.E., Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the MIG6 family.
 CC

This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; BC005546; AAH05546_1; -;

DR EMBL; BC057646; AAH57646_1; -;

DR NGD; MGI:1921405; 1300002FL3Rik.

SEQUENCER 461 MA; 50075 MW; OCA4414516FAE27A6 CRC64;
 SQ

Query Match 36.6%; Score 52; DB 1; Length 461;
 Best Local Similarity 52.9%; Pred. No. 18;

Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RPLPKPSWAPKRHRL 17
 Db 230 RVVDPDNPPQSHRL 246

Search completed: May 27, 2004, 16:19:02
 Job time : 12.5385 secs

DT	01-DEC-2001	(TREMBLrel. 19, Created)	DB	Putative E4 protein (Fragment).
DT	01-MAR-2002	(TREMBLrel. 19, Last sequence update)	GN	Human papillomavirus type 16.
DR	B4	protein (Fragment).	OS	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
GN			OC	Papillomavirus.
Human papillomavirus type 16.			OC	
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			NCBI_TaxID=10581;	
OC			RN	[1]
Papillomavirus.			RP	SEQUENCE FROM N.A.
NCBI_TaxID=10581;			RA	Terai M., Ma Z., Burk R.D.;
RN			RA	Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
RP			RT	"Sequence variation and physical state of human Papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia.";
RC	STRAIN=HPV16E2CC4;		RT	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RA	Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.,		RT	Overlapping PCR Method.;
RT	"Sequence variation and physical state of human Papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia.";		RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RL	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.		DR	InterPro: IPR003861; Papilloma_E4.
DR	EMBL; AF472508; AAO15701; -.		DR	PFam: PF02711; Pap_E4; 1.
DR	InterPro: IPR003861; AAL01396; -.		FT	NON_TER 1 1 DR
DR	Pfam; PF02711; Pap_E4; 1.		SQ	SEQUENCE 95 AA; 10585 MW; AEC1961D677307CE CRC64;
FT	NON_TER 1 1		Query Match	Best Local Similarity 85.2%; Score 121; DB 12; Length 95;
SQ	SEQUENCE 95 AA; 10608 MW; AED4269D05B307CE CRC64;		Matches	Best Local Similarity 92.6%; Pred. No. 6.7e-10; Indels 0; Gaps 2;
Query Match	85.2%; Score 121; DB 12;		Matches	Best Local Similarity 92.6%; Pred. No. 6.7e-10; Indels 0; Gaps 2;
Best Local Similarity 92.6%; Pred. No. 6.7e-10;			QY	1 RPIPKPSPWAPKKHRRRL-SDQD-SQTP 25
Matches 25; Conservative 0; Mismatches 0; Indels 2; Gaps 2;			Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RESULT 5	Q8B5N7
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		ID	Q8B5N7 PRELIMINARY; PRT; 95 AA.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		AC	Q8B5N7;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DT	01-MAR-2003 (TREMBLrel. 23, Created)
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DB	Putative E4 protein (Fragment).
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		GN	Human Papillomavirus type 16.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		OS	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		OC	Papillomavirus.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		OX	NCBI_TaxID=10581;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RN	[1]
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RP	SEQUENCE FROM N.A.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RA	Terai M., Ma Z., Burk R.D.;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RA	Cloning and Sequencing of Non-European Human Papillomavirus (HPV)
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RT	"Sequence variation and physical state of human Papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia.";
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RT	Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RL	InterPro: IPR003861; Papilloma_E4.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DR	PFam: PF02711; Pap_E4; 1.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		FT	NON_TER 1 1 DR
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		SQ	SEQUENCE 95 AA; 10542 MW; AED17903867307CE CRC64;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		Query Match	Best Local Similarity 85.2%; Score 121; DB 12; Length 95;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		Matches	Best Local Similarity 92.6%; Pred. No. 6.7e-10; Indels 0; Gaps 2;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		QY	1 RPIPKPSPWAPKKHRRRL-SDQD-SQTP 25
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		RESULT 6	Q8OMMO
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		ID	Q8OMMO PRELIMINARY; PRT; 95 AA.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		AC	Q8OMMO;
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DT	01-JUN-2003 (TREMBLrel. 24, Created)
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		DE	E4 (Fragment).
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		OS	Human Papillomavirus type 16.
Db	29 RPIPKPSPWAPKKHRRRLSDQDQSQT 55		OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.	RA Buell C.R.; Yuan Q.; Moffat K.S.; Hill J.N.; Jenkins C.N.; Hsiao J.; Zismann V.; Pai G.; Bowman C.L.; Fujii C.Y.; VanAken S.E.; Craven S.L.,
OX NCBI_TaxID=10581;	RA
RN [1]_TaxID=10581;	RA
RP SEQUENCE FROM N.A.	RA
RC STRAIN="Asian-American variant;"	RA
TERI M.; Burk R.D.;	RA
"Human papillomavirus type 16 Asian-American variant."	RT "Orziza sativa chromosome 10 BAC OSJNBA0027P10 genomic sequence."
Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.	RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF02678; AA05412; 1; -.	DR EMBL; AC084763; AAC60194; 1; -.
DR InterPro; IPR003867; Papilloma_E4.	DR Gramene; Q9AY37; GO; GO:0016305; F:kinase activity; TEA.
Pfam; PF02711; Pap_E4; 1.	DR InterPro; IPR003409; MORN.
FT NON_TER 1 1.	DR Pfam; PF02493; MORN; 7.
SQ SEQUENCE 95 AA; 10542 MW; AED17903867307GE CRC64;	DR SMART; SM00698; MORN; 7.
Query Match 85.2%; Score 121; DB 12; Length 95;	DR Kinase.
Best Local Similarity 92.6%; Pred. No. 6.7e-10;	KW
Matches 25; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	SQ SEQUENCE 467 AA; 50906 MW; 8FB85BCC9B980D73 CRC64;
Qy 1 RPIPKPSWAPKHKRLL-SQDP-SQTP 25	Query Match 47.2%; Score 67; DB 10; Length 467;
Db 29 RPIPKPSWAPKHKRLLSQDSQTP 55	Best Local Similarity 68.8%; Pred. No. 0.21;
RN	Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
RESULT 7	Qy 1 RPIPKPSWAPKHKRLL-SQDP-SQTP 16
Q99326 ID Q99326 PRELIMINARY; PRT; 88 AA.	Db 72 RPSPPSPFAPSRRR 87
AC Q99326; DT 01-JUN-2001 (TREMBLrel. 17, Created)	RESULT 9
DT 01-OCT-2001 (TREMBLrel. 17, Last sequence update)	Q7XC25 ID Q7XC25 PRELIMINARY; PRT; 467 AA.
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)	AC Q7XC25; DT 01-OCT-2003 (TREMBLrel. 25, Created)
DE Putative B4 protein.	DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
GS Human papillomavirus type 82	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;	DE Putative phosphatidylinositol-4-phosphate 5'-kinase.
OC Papillomavirus.	GN OSJNBA0027P10-22.
OX NCBI_TaxID=129724;	OS Oryza sativa (japonica cultivar-group).
RN	OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
SEQUENCE FROM N.A.	OC Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;
TERI M.; Burk R.D.;	OC Ehrhartoidae; Orzyeae; Oryza.
"Cervical HPVs in Evolution; Genomic Sequence of IS39/AE2, a Subtype	NCBI_TaxID=39947; RN [1]_TaxID=39947;
of Oncogenic HPV 82 (W13B)."	RN SEQUENCE FROM N.A.
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	RP STRAIN=CV; Nipponbare; RC
EMBL; AF23961; AAK2845; 1; -.	RA The Rice Chromosome 10 Sequencing Consortium;
InterPro; IPR003867; Papilloma_E4.	RA Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
Pfam; PF02711; Pap_E4; 1.	RT "In-depth view of structure, activity, and evolution of rice
SQ SEQUENCE 88 AA; 10084 MW; 6752D8CF2A9475D7 CRC64;	RT chromosome 10"; RL Science 300:1566-1569 (2003).
Query Match 51.1%; Score 72.5; DB 12; Length 88;	RN [2] RN SEQUENCE FROM N.A.
Best Local Similarity 50.0%; Pred. No. 0.0061;	RP STRAIN=CV; Nipponbare; RC
Matches 15; Conservative 4; Mismatches 6; Indels 5; Gaps 1;	RA Buell C.R.; Wing R.A.; McCombie W.R.; Messing J.; Yuan Q.; DR Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
Qy 1 RPIPKPSWAPKHKRLL-SQDP-SQTP 25	RA DR EMBL; AR017120; ARP55050; 1; -.
Db 24 RPIPKPSWAPKHKRLLSQDSQTP 53	KW SQ SEQUENCE 467 AA; 50906 MW; 8FB85BCC9B980D73 CRC64;
RN	Query Match 47.2%; Score 67; DB 10; Length 467;
RESULT 8	Best Local Similarity 68.8%; Pred. No. 0.21;
Q9AY37 ID Q9AY37 PRELIMINARY; PRT; 467 AA.	Matches 11; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
AC Q9AY37; DT 01-JUN-2001 (TREMBLrel. 17, Created)	Qy 1 RPIPKPSWAPKHKRLL-SQTP 16
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	Db 72 RPSPPSPFAPSRRR 87
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	RESULT 10
DE Putative phosphatidylinositol-4-phosphate 5'-kinase.	O90727 ID O90727 PRELIMINARY; PRT; 97 AA.
OS OUNBA0027P10-22.	AC O90727; DT 01-NOV-1998 (TREMBLrel. 08, Created)
OC Oryza sativa (Rice).	DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
OC Ehrhartoidae; Orzyeae; Oryza.	DE ORF E4.
OX NCBI_TaxID=4530;	OS Human papillomavirus type 67.
RN [1]_TaxID=4530;	VC viruses; no RNA stage; Papillomaviridae;
RP SEQUENCE FROM N.A.	RC
RC STRAIN=CV; Nipponbare;	

OC Papillomavirus.	Query Match Score 60;	DB 12;	Length 1242;
OX NCBI_TAXID:37120;	Best Local Similarity 42.3%;	Pred. No. 6;	
RN [1] -	Matches 11; Conservative 3;	Mismatches 9;	Indels 0;
RP SEQUENCE FROM N.A.	Gaps 0;		
RA Kirii Y., Matsuura T.;			
RT "Nucleotide sequence and phylogenetic classification of human papillomavirus type 67."			
RT Virus Genes 17:117-121(1998).			
RL EMBL; D21208; BZAA2886.1/-.			
DR InterPro; IPR000861; Papilloma_E4.			
DR Pfam; PF02711; Pap_E4; 1.			
DR Sequence 97 AA; 10881 MW; D64C424B258C888E CRC64;			
Query Match Score 64;	DB 12;	Length 97;	
Best Local Similarity 59.1%; Pred. No. 0.11;			
Matches 13; Conservative 2; Mismatches 7;	Indels 0;	Gaps 0;	
QY 3 IPKPSPWPKKHRLSDQDSQ 24			
Db 28 IPLPPPPPKCNRRLPDDDLT 49			
RESULT 11			
Q9PZW6 PRELIMINARY; PRT; 1242 AA.			
AC Q9PZW6; STRAIN=Columbia; PubMed=10907853;			
DT 01-MAY-2000 (TREMBLrel. 13, Created)			
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DE Structural polyprotein.			
OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis virus).			
CC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.			
OX NCBI_TAXID:11021;			
RP SEQUENCE FROM N.A.			
STRAIN=BR85-436087;			
RA Brault A.C., Powers A.M., Kang W., Tesh R.B., Shope R.E., Weaver S.C.,			
RT "Genetic and Antigenic Diversity among Eastern Equine encephalitis viruses from North, Central and South America.";			
RT Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.			
EMBL; AF155561; AAF04803.1; -.			
DR HSSP; P03315; 1TCP.			
DR GO; GO:0016020; C:membrane; IEA.			
DR GO; GO:0019028; C:viral envelope; IEA.			
DR GO; GO:00019031; C:viral envelope; IEA.			
DR GO; GO:00051252; F:serine-type endopeptidase activity; IEA.			
DR GO; GO:0005198; F:structural molecule activity; IEA.			
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.			
DR InterPro; IPR00548; Alpha_B1_glycop.			
DR InterPro; IPR009336; Alpha_B2_glycop.			
DR InterPro; IPR00533; Alpha_B3_glycop.			
DR InterPro; IPR009003; Cys_Ser_trypsin.			
DR InterPro; IPR000336; Flavi_glycoprote.			
DR InterPro; IPR00110; Ig-like.			
DR InterPro; IPR005006; Ig_MHC.			
DR InterPro; IPR009930; Peptidase_S3.			
DR Pfam; PE01589; Alpha_core_1.			
DR Pfam; PE00943; Alpha_B1_glycop; 1.			
DR Pfam; PE01563; Alpha_B2_glycop; 1.			
DR Pfam; PE01563; Alpha_B3_glycop; 1.			
DR PRINTS; PRO798; TOGAVIRIN.			
DR PROSITE; PS00290; IG_MHC; 1.			
FW Polypeptide.			
SEQUENCE 1242 AA; 137405 MW; 7254B6640ED65C3 CRC64;			

Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E., Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M., Houston K.A., Humasti S.R., Kim E., Li P., Moskrafi M., Pacleb J.M., Park S., Sequeira A., Sethi H., Snir E., Svirskaas R.R., Weinburg T., Celinker S.E.; "Full length *Drosophila melanogaster* cDNA sequence."; Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

DR FlyBase; FBgn0275585; BCDN:GH05582.

SQ Sequence 985 AA; 112852 MW; 6687651C050BF4F CRC64;

Query Match ID Q9VAXX8 PRELIMINARY; PRT; 985 AA.

AC Q9VAXX8; PRELIMINARY; PRT; 985 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAR-2003 (TREMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE CG8740 protein. BCNA:GH0582 OR CG8740.

OS *Drosophila melanogaster* (Fruit fly).

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophila; *Drosophila*.

OC NCBI_TaxID=7227;

OX STRAIN=Berkely;

RP SEQUENCE FROM N.A.

RX Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scheer S.E., Li P.W., Hoskins R.A., Galie R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.H.C., Blaej R.G., Champé M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Heit G., Neilson C.R., Milios G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D., Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M., Benson K.Y., Benos P.V., Bernier B.P., Bhandari D., Boishartov S., Borikova D., Botchan M.R., Bouck J., Brookstein P., Brottier P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablo B., Delcher A., Deng Z., Dew I., Diaz S.M., Dodge K., Doucet L.F., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin R., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foster C., Gabrilian A.E., Ferraz C., Glasser K., Harris M., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Hattori M.L., Harvey D., Heiman-T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Hernaei M.-H., Ikegami C., Jaijali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moskrafi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pan G.S., Pan S., Pollard J., Puris V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shiu B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Turner E., Venter J.C., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodage T., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhou W., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of *Drosophila melanogaster*."; Science 287:2185-2195 (2000). [2]

RP SEQUENCE FROM N.A.

RA Celinker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y., Barzon J., An H., Baldwin D., Banzon J., Beson K.Y., Busam D.A., Carlson K., Dorsett V., Center L., Champé M., Davenport L.B., Dietz S.M., Dodson K., Dorschett V., Doup L.E., Doyle C., Dresneak D., Farfan D., Ferrier S., Frise E., Galie R.F., Garg N.S., George R.A., Gonzalez M., Houck J., Hoskins R., Hostin D., Howland T.J., Ibewalew C., Jalali M., Kruse D., Li P., Mattei B., Moskrafi A., McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J., Pacek J., Paragas V., Park S., Patel S., Pfeiffer B., Puris V., Richards G.S., Pittman G.S., Puris V., Scheeler F., Williams S.M., Zaveri J.S., Zhan M., Zhang G., Zhou W., Zhu S., Zhu X., Smith H.O., Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases. [3]

RN SEQUENCE FROM N.A.

RA Misa S., Crosby M.A., Mattbews B.B., Bayraktaroglu L., Campbell K., Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D., Tupy J.L., Bergman C., Berman B., Carlson J.W., Celinker S.E., Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N., Krommiller B., Marshall B., Millburn J., Richter J., Russo S., Sealje S.M.J., Smith E., Shu S., Smutniak F., Whitfield E., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.; "Annotation of *Drosophila melanogaster* genome."; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases. [4]

RN SEQUENCE FROM N.A.

RA Adams M.D., Celinker S.E., Gibbs R.A., Rubin G.M., Venter C.J.; Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases. [5]

RP SEQUENCE FROM N.A.

RA Flybase; Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases. [5]

RN DR EMBL; AB013836; ADM68816.1; -

DR FIBNC0027585; BODNA-GH05582.

SQ SEQUENCE 985 AA; 112894 MW; 245F1B903DAC636 CRC64;

Query Match Score 58; DB 5; Length 985; Best Local Similarity 52.4%; Matches 11; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

QY 1 RPIPKPSWPKKHERLSDDD 21

Db 106 RPSHSPSPWQQLSPRNLSDD 126

RESULT 15

Q9TR55 PRELIMINARY; PRT; 88 AA.

AC Q9TR55; PRELIMINARY; PRT; 88 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

GN E4.

RA Human papillomavirus type 82. Viruses, dsDNA viruses, no RNA stage, Papillomaviridae, Papillomavirus.

OC OC

RA Sata T., Matsukura T.; Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases. [2]

RN RN

RP SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RN SEQUENCE FROM N.A.

RX MEDLINE-20087389; PubMed=10618284;

RX Kino N., Sata T., Sato Y., Matsukura T.;

RT "Molecular cloning and nucleotide sequence analysis of a novel human papillomavirus (type 82) associated with vaginal intraepithelial neoplasia";
 RT Clin. Diagn. Lab. Immunol. 7:91-95 (2000).
 RL EMBL; AB027021; BAA90739.1; -
 DR IPR003861; Papilloma_E4.
 DR InterPro; PF02711; Pap_E4_1.
 DR PFAM; PF02711; Pap_E4_1.
 SQ SEQUENCE 88 AA; 9960 MN; 9776FA242AE3202B CRC64;

Query Match	Score	Length	Gaps
Best Local Similarity	40.5%	88	
Matches	46.4%	12	
Conservative	46.4%	88	
Mismatches	3	3	
Indels	9	3	

QY 1 RPPKPSFWAPKK--HRRLLSDQSQTQ 25

Search completed: May 27, 2004, 16:20:55
Job time : 50.0769 secs

Copyright (c) 1993 - 2004 Compugen Ltd.	GenCore version 5.1.6	Aab73488 Mouse Rim			
OM protein - protein search, using sw model		Aau48473 Propionib			
Run on: May 27, 2004, 16:03:03 ; Search time 41.5385 Seconds	(without alignments)	Aab44992 Propionib			
Perfect score: 91	102.031 Million cell updates/sec	Aab23168 Human bra			
Sequence: 1 RPIPKPSPWAPKKHR 15		Aab23167 Human bra			
Title: US-10-008-524A-167		Aab23167 Human bra			
Scoring table: BLOSUM62		Abrg13930 Novel hum			
Gapop 10.0 , Gapext 0.5		Aao00838 Human pol			
Searched: 1586107 seqs, 282547505 residues		Aab23160 Bucalyptu			
Total number of hits satisfying chosen parameters:	1586107	Aab17524 Novel hum			
Minimum DB seq length: 0		Aau59082 Propionib			
Maximum DB seq length: 2000000000		Abm5501 Propionib			
Post-processing: Minimum Match 0%	Maximum Match 100%	Abrg0584 Novel hum			
	Listing first 45 summaries	Aay59694 Secreted			
Database :	A_Geneseq_29Jan04:*	Aay36143 Human sec			
	1: geneseqp1980s:*	Aay36190 Human sec			
	2: geneseqp1990s:*	Aau64615 Propionib			
	3: geneseqp2000s:*	Abm61134 Propionib			
	4: geneseqp2001s:*	Abr41361 Human DIT			
	5: geneseqp2002s:*	Aag01947 Human sec			
	6: geneseqp2003as:*				
	7: geneseqp2003bs:*				
	8: geneseqp2004s:*				
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
		SUMMARIES			
Result No.	Score	Query Match	Length	DB ID	Description
-	-	-	-	-	-
1	91	100.0	26	2 AAW62277	Aaw62277 HPV16 hyd
2	91	100.0	26	2 AAU75220	Aau7260 Hydrophil
3	91	100.0	30	2 AAR15564	Aar15564 Immunopep
4	82	90.1	15	2 AAW62278	Aaw62278 HPV16 hyd
5	79	85.8	12	2 AAR15562	Aar15562 Immunopep
6	75	82.4	12	2 AAW62279	Aaw62279 HPV16 hyd
7	68	74.7	14	2 AAR14288	Aar14288 Seroreact
8	66	72.5	11	2 AAR14287	Aar14287 Seroreact
9	56	61.5	105	4 AAO03023	Aao03023 Human pol
10	54	59.3	1053	7 ADE54638	Ades54638 Human Pro
11	54	59.3	1553	7 ADE54636	Ades54636 Rat Prote
12	51	56.0	135	4 AAU48618	Aau48618 Propionib
13	51	56.0	135	6 ABM45137	Abm45137 Propionib
14	50.5	55.5	137	4 ABB11595	Abb11595 Human FPM
15	50.5	55.5	967	8 ADE52670	Ades52670 Human pro
16	50	54.9	55	4 ABG16750	Abg16750 Novel hum
17	50	54.9	115	4 AA004150	Aa004150 Human pol
18	49	53.8	58	4 AAU41082	Aau41082 Propionib
19	49	53.8	58	6 ABM37601	Abm37601 Propionib
20	49	53.8	142	2 AAY48286	Aay48286 Human pro
21	48.5	53.3	1514	6 ABJ19753	Abj19753 Human MP2
22	48	52.7	123	6 ABP96561	Abp96561 Human pro
23	48	52.7	177	7 ABM74128	Abm74128 DNA clone
24	48	52.7	273	6 ABU41964	Abu41964 Protein e
25	48	52.7	594	4 ABB61362	Abb61362 Drosophil

ALIGNMENTS

RESULT 1
AAW62277 standard; peptide; 26 AA.
ID AAW62277
XX AC AAW62277;
XX DT 24-SEP-1998 (first entry)
XX DB HPV16 hydrophilic region peptide #1.
XX KW HPV16; human papilloma virus; epithelial tumour; cervical cancer;
KW precancerous cervical lesion; screening; detection; infection; cervix;
KW HPV E4.
XX OS Human papillomavirus.
XX PN WO9825145-A1.
XX PD 11-JUN-1998.
XX XX 03-DEC-1997; 97WO-GB003321.
XX PR 03-DEC-1996; 96GB-00025142.
XX PR 05-SEP-1997; 97GB-00018745.
XX PA (MEDI) MEDICAL RES COUNCIL.
XX PI Doorbar J;
XX DR WPI; 1998-333497/29.
XX PT Detecting papilloma virus infection using molecule binding to E4 protein - useful, e.g. in screening for pre-cancerous cervical lesions and to determine type(s) of human papilloma virus infecting human patients.
XX PS Claim 9; Page 37; 52pp; English.
XX CC A new method has been developed for detecting a papilloma virus infection in an organism. The method comprises: (i) obtaining a sample of cells from the potential infection site; (ii) contacting the cells with a molecule binding specifically to papilloma virus E4 protein, and (iii) monitoring the binding. The method is useful to detect papilloma virus infections in organisms (e.g. mammals) and especially HPV infections (e.g. with HPV16, 18, 35, 45, 51, 56, 58 or 61) in humans. Papilloma viruses cause epithelial tumours in humans varying in severity depending on the infection site and HPV type involved. The method is particularly useful to determine papilloma infection in the mammalian cervix and especially to screen for pre-cancerous cervical lesions in

CC humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in CC a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and hence clinical prognosis) is dependent on HPV type. Molecules capable of CC binding E4 are also useful to target anticancer/antiviral agents capable of destroying papilloma viruses and/or papilloma virus-infected cells. CC The present sequence represents a specifically claimed HPV16 peptide CC found in the hydrophilic region

XX Sequence 26 AA;

Query Match 100.0%; Score 91; DB 2; Length 26;

Best Local Similarity 100.0%; Pred. No. 2.2e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPPIPKPSPWAPKKHR 15

Db 1 RPPIPKPSPWAPKXHR 15

RESULT 2
AAU75260 standard; peptide; 26 AA.

XX ID AAU75260 standard; peptide; 26 AA.

XX AC AAU75260;

XX DT 21-MAY-2002 (first entry)

XX DE Hydrophilic region of HPV16 E4 protein.

XX KW Papilloma virus associated antigen; human papillomavirus infection; HPV; lesion;

XX KW cervical malignancy; human papillomavirus marker; HPV; lesion;

XX KW cellular abnormality; cellular proliferation; cellular growth; dysplasia;

XX KW neoplasia; cancer; papilloma smear test; E4 protein.

XX OS Human papillomavirus type 16.

XX FN WO200208764-A1.

XX PR 31-JAN-2002.

XX PT 16-MAR-2001; 2001WO-GB001176.

XX BR 24-JUL-2000; 2000GB-00018140.

XX PA (MEDI-) MEDICAL RES COUNCIL.

XX PI Doorbar Ü;

XX DR WPI; 1992-188648/24.

XX Disclosure; Page 23; 90pp; English.

XX Detecting abnormalities e.g. abnormal cellular proliferation, in a sample PT from a patient comprises contacting cells with a molecule which can bind PT a papilloma virus associated antigen, or a cell proliferation or viral PT activity marker.

XX Disclosure; Page 23; 90pp; English.

XX The present invention relates to a method for detecting abnormalities in CC a sample from a patient. The method comprises contacting a sample of the CC patient's cells with two or more molecules, where at least one molecule CC is capable of binding a papilloma virus associated antigen, and at least one molecule is capable of binding a cell proliferation marker. The CC method is useful for simultaneously screening for abnormalities which indicate or can lead to cervical malignancy, for human Papillomavirus CC (HPV) infections, and precursor lesions or other conditions which occur CC with cervical malignancy. The method is also useful for assessing the CC risk associated with cellular abnormality in a patient sample, and for determining, assessing or diagnosing the presence or absence of abnormal CC cellular proliferation, cellular growth abnormality, dysplasia, CC neoplasia, or a pre-cancerous or cancerous state in a tissue. The new CC method is much simpler, and yields more information more quickly than

CC conventional papilloma smear testing programmes. Compared with previous CC methods of screening, the new method has reduced chances of false CC negatives occurring, requires fewer samples to gain the same amount of CC information, and alleviates the need for repeated or further testing. The CC present sequence representing the hydrophilic region of the HPV16 E4 CC protein binds antibody molecules

XX SQ Sequence 26 AA;

Query Match 100.0%; Score 91; DB 5; Length 26;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPPIPKPSPWAPKXHR 15

Db 1 RPPIPKPSPWAPKXHR 15

RESULT 3
AAR15564 standard; protein; 30 AA.
ID AAR15564
XX AC AAR15564;
XX DT 02-MAR-1992 (first entry)
XX DE Immunopeptide #4 derived from HPV16 E4 Peptide
XX KW cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
KW squamous cell carcinoma; ELISA; HPV 16.
XX OS Synthetic.
XX PN WO9118294-A.
XX PD 28-NOV-1991.
XX PF 11-MAY-1990; 90SE-00001705.
XX PR 11-MAY-1990; 90SE-00001705.
XX PA (MEDS-) MEDSCAND AB.
XX PI Dillner J, Dillner L, Cheng HM,
XX DR WPI; 1991-369390/50.
XX PT Diagnosis of human Papilloma virus infection and PV-carrying tumours -
PT using synthetic peptide(s) to detect virus specific antigen-antibody
XX complexes by immunoassay.
XX Disclosure; Page 38; 72pp; English.

CC This is one of a large number of peptides which have been synthesised on CC the basis of the amino acid sequences for the E2, E4, E7, L1 or L2 CC proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The selection of CC peptide sequences was based on the assumption that an immunoreactive CC region might be situated in the same relative region of a protein from CC different HPV types. The peptides were used in diagnostic immunoassays to CC detect HPV-infection. See AAR15523-R15601

XX SQ Sequence 30 AA;

Query Match 100.0%; Score 91; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPPIPKPSPWAPKXHR 15

Db 7 RPPIPKPSPWAPKXHR 21

RESULT 4

XX	(MEDI-) MEDICAL RES COUNCIL.	PT Sero-active epitope(s) of human papilloma-virus 16 proteins - for use vaccines and diagnosis.
PA		PT
XX		XX
PI	Doorbar J;	PS Claim 1; Page 11; 15pp; English.
XX		XX
DR	WPI; 1998-333497/29.	CC This is one of two seroreactive epitopes identified from HPV16 protein E4. Peptides containing this epitope are also claimed and can be used to generate antibodies to HPV. See AAR14287-R14302 and AAQ14168-Q14171.
XX	Detecting papilloma virus infection using molecule binding to E4 protein - useful, e.g. in screening for pre-cancerous cervical lesions and to determine type(s) of human papilloma virus infecting human patients.	CC
PR		CC
PT		CC
PP		CC
XX		XX
ES	Claim 11; Page 37; 52pp; English.	SQ Sequence 14 AA;
XX		Query Match 74.7%; Score 68; DB 2; Length 14;
CC	A new method has been developed for detecting a papilloma virus infection in an organism. The method comprises: (i) obtaining a sample of cells from the potential infection site; (ii) contacting the cells with a molecule binding specifically to papilloma virus E4 protein, and (iii) monitoring the binding. The method is useful to detect papilloma virus infections in organisms (especially mammals) and especially HPV infections (e.g. with HPV16, 18, 31, 33, 45, 51, 56, 58 or 61) in humans. Papilloma viruses cause epithelial tumours in humans, varying in severity depending on the infection site and HPV type involved. The method is particularly useful to determine papilloma infection in the mammalian cervix and especially to screen for pre-cancerous cervical lesions in humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and hence clinical prognosis) is dependent on HPV type. Molecules capable of binding E4 are also useful to target anticancer/antiviral agents capable of destroying papilloma viruses and/or papilloma virus-infected cells. The present sequence represents specifically claimed HPV16 peptide found in the hydrophilic region	Best Local Similarity 100.0%; Pred. No. 0.0014; Indels 0; Gaps 0;
CC		Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC		Qy 5 KPSWPWPKHR 15
CC		Db 1 KPSWPWPKHR 11
CC		XX
CC		RESULT 8
CC		AAR14287
CC		ID AAR14287 standard; peptide; 11 AA.
CC		XX
CC		AC AAR14287;
CC		XX
CC		DT 02-JAN-1992 (first entry)
CC		XX
CC		DE Seroreactive epitope #1 of HPV16 protein E4.
CC		XX
CC		DE HPV16-dependent human disease; E6; E7; L1.
CC		XX
CC		OS Synthetic.
CC		XX
CC		PN EP451550-A.
CC		XX
CC		PD 16-OCT-1991.
CC		XX
CC		PF 19-MAR-1991; 91EP-00104197.
CC		XX
CC		PR 20-MAR-1990; 90EP-00105222.
CC		XX
CC		PA (BEHW) BEHRINGWERKE AG.
CC		XX
CC		PI Muller M, Giessmann L;
CC		XX
CC		DR WPI; 1991-304643/42.
CC		XX
CC		PT Sero-active epitope(s) of human papilloma-virus 16 proteins - for use vaccines and diagnosis.
CC		XX
CC		PS Claim 1; Page 11; 15pp; English.
CC		XX
CC		DE Seroreactive epitope #2 of HPV16 protein E4.
CC		XX
CC		DE HPV16-dependent human disease; E6; E7; L1.
CC		XX
CC		OS Synthetic.
CC		XX
CC		PN EP451550-A.
CC		XX
CC		PD 16-Oct-1991.
CC		XX
CC		PF 19-MAR-1991; 91EP-00104197.
CC		XX
CC		PR 20-MAR-1990; 90EP-00105222.
CC		XX
CC		PA (BEHW) BEHRINGWERKE AG.
CC		XX
CC		Muller M, Giessmann L;
CC		XX
CC		DR WPI; 1991-304643/42.
CC		XX
CC		PT Sero-active epitope #2 of HPV16 protein E4.
CC		XX
CC		DE Seroreactive epitope #2 of HPV16 protein E4.
CC		XX
CC		DE HPV16-dependent human disease; E6; E7; L1.
CC		XX
CC		OS Synthetic.
CC		XX
CC		Qy 3 IPKPSWPWPKK 13
CC		Db 1 IPKPSWPWPKK 11
CC		XX
CC		RESULT 9
CC		AAO03023
CC		ID AAO03023 standard; protein; 105 AA.
CC		XX
CC		AC AAO03023;

XX XX 06-NOV-2001 (first entry)
 DT XX Human polypeptide SEQ ID NO 16915.
 XX DE
 XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX OS Homo sapiens.
 PN WC200164835-A2.
 XX PD 07-SEP-2001.
 XX PF 26-FEB-2001; 2001WO-US004927.
 XX PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX PA (HYSB-) HYSBQ INC.
 PI Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-514838/56.
 DR N-PDB; AAI82954.
 PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 XX PS Claim 20; SEQ ID NO 16915; 1399pp + Sequence Listing; English.
 CC The invention relates to human polynucleotides (AAI79941-AAI9381) and
 CC the encoded proteins (AAQ0010-AAQ13910) that exhibit activity relating to
 CC cell proliferation or cell differentiation which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haemopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp://wipo.int/pub/published_pct_sequences
 XX SQ Sequence 105 AA;
 CC Query Match 61.5%; Score 56; DB 4; Length 105;
 CC Best Local Similarity 81.8%; Pred. No. 4.1;
 CC Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 CC
 CC QY 4 PKPSPWAKKH 14
 CC | ||||| | | |
 CC 82 PCPSPWAKRH 92
 CC
 CC RESULT 1.0
 CC ADE54638
 CC ID ADE54638 standard; protein; 1053 AA.
 AC ADE54638;
 XX DT 29-JAN-2004 (first entry)
 XX ID Human Protein BAA20798, SEQ ID NO 443.
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 OS Homo sapiens.

XX PN WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PF 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-031147P.
 PR 01-NOV-2001; 2001US-034382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX P1 Woolf C, D'urso D, Befort K, Costigan M;
 XX DR WPI; 2003-268812/26.
 DR GENBANK; BAA22798.
 XX PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX PS Claim 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variant of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition,
 CC a method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides and their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp://wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 1053 AA;
 CC Query Match 59.3%; Score 54; DB 7; Length 1053;
 CC Best Local Similarity 53.8%; Pred. No. 73;
 CC Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 2 PIPKPSWPWAKKH 14
 CC | : | : | : | : | : | : | : |
 CC Db 890 PLPQQSPMPRRH 902
 CC
 CC RESULT 1.1
 CC ADE54636
 CC ID ADE54636 standard; protein; 1553 AA.
 AC ADE54636;
 XX DT 29-JAN-2004 (first entry)
 XX DE Human Protein BAA20798, SEQ ID NO 443.
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 OS Homo sapiens.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury; chronic constriction injury; CCI; spared nerve injury; SNI; Chung, XX
 Rattus norvegicus.
 OS
 XX
 WO2003016475-A2.
 PD
 27-FEBB-2003.
 XX
 PR 14-AUG-2002; 2002WO-US025165.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333337P.
 XX
 PA (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR GENBANK; AAB66703.
 XX
 PR New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
 PT
 XX
 PS Claim 1; Page: 1017pp; English.
 XX
 The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also, claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subject to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotide, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 Sequence 1553 AA;

Query Match 59.3%; Score 54; DB 7; Length 1553;
 Best Local Similarity 53.8%; P-Red. No. 1.1e+02;
 Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PIPKPSWAPKKH 14
 Db 890 PLQQPSPEMRRH 902

RESULT 12
 AAU48618
 ID AAU48618 standard; protein; 135 AA.
 XX
 AC AAU48618;
 XX

DT 27-FEB-2002 (first entry)
 XX Propionibacterium acnes immunogenic protein #9514.
 DE XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteoprotectant.
 KW XX
 OS Propionibacterium acnes.
 XX
 PN WO00181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PP 20-APR-2001; 2001WO-US012865.
 XX
 PR 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Seikly YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR 2001-616774/71.
 DR N-PSDB; AAS59543.
 XX
 Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris.
 XX
 PS Example 1; SEQ ID NO 9813; 1069p; English.
 XX
 Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to downregulate expression and activity of P. acnes polypeptides and therefore treat P. acnes infections. The antibodies may also be used as diagnostic agents for determining P. acnes presence. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 Sequence 135 AA;

Query Match 56.0%; Score 51; DB 4; Length 135;
 Best Local Similarity 66.7%; Pred. No. 24;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPIPKPSWAPKK 12
 Db 69 RPPIPMSSWPPR 80

RESULT 13
 ABM45137
 ID ABM45137 standard; protein; 135 AA.
 XX
 AC ABM45137;
 XX

DT	20-OCT-2003	(First entry)	ID	ABB11595 standard; peptide; 137 AA.
XX			XX	ABB11595;
DE			AC	
XX			XX	
XW	Propionibacterium acnes predicted ORF-encoded polypeptide #9813.		DT	11-JAN-2002 (first entry)
XX	Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;		XX	
XX	immunostimulant; immune response; vaccine.		XX	Human FPM315 homologue, SEQ ID NO:1965.
OS	Propionibacterium acnes.		XX	
XX	WO2013033315-A1.		XX	Human; cytokine; cell proliferation; cell differentiation; growth factor; haematopoiesis regulation; tissue growth; immunomodulator; onogenesis; inhibin; chemotaxis; chomokinesis; thrombolytic; oncogenesis; proliferation; metastasis; cancer; tumour; haemopoietic disorder; myeloid cell disorder; lymphoid cell disorder; asthma; arthritis; chronic inflammatory condition; proliferative retinopathy; atherosclerosis; coronary heart disease; arterial ischaemia; bone disorder; osteoporosis; vascular growth disorder; tissue regeneration; wound healing; infection; immune disorder; antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic; cytostatic; osteopathic; vasoergic; cardiot; viricide; antibacterial; antifungal; pulmonary; antiulcer.
PN	WO2013033315-A1.		XX	
XX	24-APR-2003.		XX	
PD			XX	
XX	11-OCT-2002; 2002WO-US032727.		XX	
XX	PR 15-OCT-2001; 2001US-00978825.		XX	
XX	(CORTI-) CORIXA CORP.		XX	
PI	Mitcham JL, Skeiky YAW, Persing DH, Bhartia A, Maisonneuve JL; Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D; Barth B, Vallieve-Douglass J;		XX	
PS	WPI: 2003-381789/36.		XX	
DR	N-PSDB; AF64472.		XX	
XX	DR WO2003-381789/36.		XX	
PT	New Propionibacterium acnes polypeptides and polynucleotides encoding the polypeptide, useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein.		XX	
PT	Example 1; SEQ ID NO 9813; 1481pp; English.		XX	
CC	The invention relates to an isolated polynucleotide (ACP64435-ACP64733) encoding a Propionibacterium acnes protein. The invention also relates to polypeptides encoded by the polynucleotides (ABA135624-ABM64536) and to immunogenic fragments of P. acnes polypeptides. The invention additionally encompasses expression vectors and host cells comprising a polynucleotide of the invention; antibodies against polypeptides of the invention; fusion proteins comprising a polypeptide of the invention; a method for stimulating an immune response specific for a P. acnes polypeptide and an isolated T cell population comprising T cells prepared via this method; a vaccine composition (comprising P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations, or antigen-presenting cells that express the polypeptide); a method and kit for detecting or determining the presence or absence of P. acnes in a patient; and a method for inhibiting the development of P. acnes in a patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion proteins, T cell populations or antigen-presenting cells that express the polypeptides are useful for diagnosing, preventing or treating acne vulgaris, or for stimulating an immune response specific for a P. acnes protein. The polynucleotides can also be used as probes or primers for nucleic acid hybridisation. The vaccine composition is useful for the stimulation of an immune response against P. acnes, or for treating acne, and the kit is useful for performing a diagnostic assay. The present sequence represents a polypeptide predicted to be encoded by an ORF (open reading frame) contained within the P. acnes polynucleotides of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences		XX	
CC	Sequence 135 AA;		XX	
CC	Query Match 56.0%; Score 51; DB 6; Length 135;		XX	
CC	Best Local Similarity 66.7%; Pred. No. 24; Mismatches 1; Indels 0; Gaps 0;		CC	
QY	1 RP1PKPSPNAPK 12		CC	
Db	69 RP1PMPPSPNPR 80		CC	
RESULT 14	ABB11595		CC	

CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.

SQ Sequence 137 AA:

Qy	1 RPIPK--PSPWAP 11	Score 50.5;	DB 4;	Length 137;
Matches 9;	Conservative	Pred. No. 29;	Mismatches 1;	Indels 3;
Db	96 RPLPKNARPSWPWP 109		Gaps 1;	

RESULT 15

ADE52670

ID ADE52670 standard; protein; 967 AA.

XX

AC ADE52670;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human protein SEQ ID 36.

XX

KW Human; DNA-binding protein; interferon-activatable protein.

XX

Homo sapiens.

XX

OS WO2003089466-A1.

XX

BN

XX

PD 30-OCT-2003.

XX

PF 18-APR-2003; 2003WO-JP004981.

XX

PR 19-APR-2002; 2002JP-00117840.

PR 30-APR-2002; 2002JP-00128418.

PR 30-APR-2002; 2002JP-00128779.

PR 04-DEC-2002; 2002JP-00352469.

XX

PA (RIKE) RIKEN KK.

PA (DNAF-) DNAFORM KK.

PA (MITU) MITSUBISHI CHEM CORP.

XX

PI Hayashizaki Y, Kamiya M, Kubodera H;

XX

WPI; 2004-011681/01.

DR N-PSDD; ADE52666.

XX

PT Proteins with DNA binding activity and substances that affect their

PT activity or expression, useful for treating associated disorders.

XX

BS Claim 1: SEQ ID NO 36; 237pp; Japanese.

XX

CC The present invention relates to novel proteins (ADE52648-ADE52660,

CC ADE5270 and ADE52672) and their coding sequences (ADE52635-ADE52647,

CC ADE52669 and ADE52621). The proteins have a DNA-binding

CC activity or expression, useful for treating associated disorders.

XX

SQ Sequence 967 AA:

Qy	Query Match 1 RPIPK--PSPWAP 11	Score 50.5;	DB 8;	Length 967;
Best Local Similarity 9;	Conservative	Pred. No. 2e+02;	Mismatches 1;	Indels 3;
Matches 9;		Gaps 1;		

Db 190 RPLPKNARPSWPWP 203

Search completed: May 27, 2004, 16:18:29
 Job time : 43.5385 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
 Run on: May 27, 2004, 16:15:59 ; Search time 12.1154 Seconds
 (without alignments)
 63.918 Million cell updates/sec

Title: US-10-008-524A-167
 Perfect score: 91
 Sequence: 1. RPIPKPSPWAPKHR 15

Scoring table: BL0SUM62
 Gapext 0.5

Searched: 383414 seqs, 51625371 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents AA:
 1: /cgn2_6/_ptodata/2/1aa/5A_COMB.pep:
 2: /cgn2_6/_ptodata/2/1aa/5B_COMB.pep:
 3: /cgn2_6/_ptodata/2/1aa/5A_COMB.pep:
 4: /cgn2_6/_ptodata/2/1aa/5B_COMB.pep:
 5: /cgn2_6/_ptodata/2/1aa/PCUTS_COMB.pep:
 6: /cgn2_6/_ptodata/2/1aa/backfiles1.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	91	100.0	15	4	US-09-314-268-167	Sequence 167, App
2	91	100.0	25	4	US-09-314-268-4	Sequence 4, App
3	91	100.0	30	2	US-08-934-915-42	Sequence 42, App
4	91	100.0	92	4	US-09-314-268-90	Sequence 90, App
5	79	86.8	20	2	US-08-934-915-40	Sequence 40, App
6	79	86.8	20	2	US-08-934-915-170	Sequence 170, App
7	75	62.4	12	4	US-09-314-268-117	Sequence 168, App
8	68	74.7	35	4	US-09-314-268-117	Sequence 117, App
9	54	59.3	173	4	US-09-252-991A-32711	Sequence 32711, A
10	52	57.1	8	4	US-09-314-268-31	Sequence 31, App
11	52	57.1	8	4	US-09-314-268-33	Sequence 33, App
12	52	57.1	8	4	US-09-314-268-37	Sequence 37, App
13	51	56.0	8	4	US-09-314-268-36	Sequence 36, App
14	51	56.0	34	4	US-09-314-268-118	Sequence 118, App
15	50	54.9	8	4	US-09-314-268-34	Sequence 34, App
16	50	54.9	8	4	US-09-314-268-35	Sequence 35, App
17	50	54.9	195	4	US-09-252-991A-22162	Sequence 22162, A
18	49	53.8	8	4	US-09-314-268-32	Sequence 32, App
19	49	53.8	325	4	US-09-921-099A-21	Sequence 21, App
20	49	53.8	419	4	US-09-252-991A-22888	Sequence 22888, A
21	48	52.7	143	4	US-09-252-991A-21367	Sequence 21367, A
22	47	51.6	154	4	US-09-252-991A-5612	Sequence 25612, A
23	46	50.5	8	4	US-09-314-268-30	Sequence 30, App
24	46	50.5	81	4	US-09-247-155-125	Sequence 125, App
25	46	50.5	81	4	US-09-663-600A-101	Sequence 101, App
26	46	50.5	81	4	US-09-663-600A-195	Sequence 195, App
27	46	50.5	200	1	US-08-187-829-3	Sequence 3, App

ALIGNMENTS

```

RESULT 1
US-09-314-268-167
; Sequence 167, Application US/09314268
; Patent No. 6346377
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314, 268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314, 268
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 167
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-314-268-167

```

```

Query Match 100.0%; Score 91; DB 4; Length 15;
Best Local Similarity 100.0%; Score 91; DB 4; Length 15;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 2
US-09-314-268-4
; Sequence 4, Application US/09314268
; Patent No. 6346377
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 3789/80902
; CURRENT APPLICATION NUMBER: US/09/314, 268
; CURRENT FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/314, 268
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-314-268-4

```

Query Match 100.0%; Score 91; DB 4; Length 25;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPIPKPSWAPKKHR 15
 Db 1 RPIPKPSWAPKXHR 15

RESULT 3
 US-08-934-915-42
 ; Sequence 42, Application US/08934915
 ; GENERAL INFORMATION:
 ; APPLICANT: DILLNER, JOAKIM
 ; APPLICANT: DILLNER, LENA
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
 ; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6,
 ; TITLE OF INVENTION: 11, 16, 18, 31, 33 AND 56,
 ; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
 ; TITLE OF INVENTION: DIAGNOSTIC PURPOSES
 ; NUMBER OF SEQUENCES: 193
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MASON & ASSOCIATES, P.A.
 ; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
 ; CITY: CLEARWATER
 ; STATE: FLORIDA
 ; COUNTRY: U.S.A.
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: Windows 3.0
 ; SOFTWARE: Microsoft Word 6.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/934,915
 ; FILING DATE: 22-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: LOUISE A. FOUTCH
 ; REGISTRATION NUMBER: 37,133
 ; REFERENCE/DOCKET NUMBER: 1946.6
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 813-538-3800
 ; TELEFAX: 813-538-3820
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 42:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 30 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Peptide
 ; US-08-934-915-42

Query Match 100.0%; Score 91; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 3.7e-06; Mismatches 0; Indels 0; Gaps 0;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPIPKPSWAPKXHR 15
 Db 7 RPIPKPSWAPKXHR 21

RESULT 4
 US-09-314-268-90
 ; Sequence 90, Application US/09314268
 ; Patent No. 6346377
 ; GENERAL INFORMATION:
 ; APPLICANT: Doorbar, John

RESULT 6
 US-08-934-915-170
 Sequence 170, Application US/08934915
 Patent No. 5932412

GENERAL INFORMATION:
 APPLICANT: DILLNER, JOAKIM
 APPLICANT: DILLNER, LENA
 APPLICANT: CHENG, HWEE MING
 TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
 PAPILLOMAVIRUS 1, 5, 6, 8,
 11, 16, 18, 31, 33 AND 56,
 USEFUL IN IMMUNOASSAY FOR
 DIAGNOSTIC PURPOSES
 NUMBER OF SEQUENCES: 193

CORRESPONDENCE ADDRESS:
 ADDRESSEE: MASON & ASSOCIATES, P.A.
 STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
 CITY: CLEARWATER
 STATE: FLORIDA
 COUNTRY: U.S.A.

COMPUTER READABLE FORM:
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: Windows 3.0
 SOFTWARE: Microsoft Word 6.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/934,915
 FILING DATE: 22-SEP-1997
 CLASSIFICATION: 435
 PRIORITY DATA:
 PRIOR APPLICATION NUMBER: 07/949,836
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: LOUISE A. FOUTCH
 REGISTRATION NUMBER: 37,133
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 813-538-3800
 TELEX:
 INFORMATION FOR SEQ ID NO: 170:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide

US-08-934-915-170
 Query Match Score 79%; DB 2; Len 168
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0;

QY 3 IPKPSWPARKHR 15
 DBb 1 IPKPSWPARKHR 13

RESULT 7
 US-09-314-268-168
 Sequence 168, Application US/09314268
 Patent No. 6345377

GENERAL INFORMATION:
 APPLICANT: DOOBAR, JOHN
 TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO S
 TITLE OF INVENTION: VIRUSES

QY 3 IPKPSWPARKHR 15
 DBb 1 IPKPSWPARKHR 13

Query Match 59.3%; Score 54; DB 4; Length 173;
 Best Local Similarity 72.7%; Pred. No. 2 3;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 SEQ ID NO: 314-268-31

Db 1 PTPKPSPWAP 11
 Db 158 RPVRISWPWAP 168

RESULT 12
 US-09-314-268-37
 ; Sequence 31, Application US/09314268
 ; Patent No. 6346377
 ; GENERAL INFORMATION:
 ; APPLICANT: Doorbar, John
 ; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
 ; FILE REFERENCE: 3789/80902
 ; CURRENT APPLICATION NUMBER: US/09/314,268
 ; CURRENT FILING DATE: 1999-03-19
 ; EARLIER APPLICATION NUMBER: 09/314,268
 ; EARLIER FILING DATE: 1999-05-18
 ; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 37
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: VIRUSES
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: octapeptide antigen
 ; SEQ ID NO: 31
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: octapeptide antigen
 ; SEQ ID NO: 314-268-31

Query Match 57.1%; Score 52; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO: 314-268-31

Db 2 PIPKPSW 9
 Db 1 PIPKPSW 8

RESULT 13
 US-09-314-268-36
 ; Sequence 36, Application US/09314268
 ; Patent No. 6346377
 ; GENERAL INFORMATION:
 ; APPLICANT: Doorbar, John
 ; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
 ; FILE REFERENCE: 3789/80902
 ; CURRENT APPLICATION NUMBER: US/09/314,268
 ; CURRENT FILING DATE: 1999-03-19
 ; EARLIER APPLICATION NUMBER: 09/314,268
 ; EARLIER FILING DATE: 1999-05-18
 ; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 36
 ; LENGTH: 8
 ; TYPE: PRT
 ; ORGANISM: VIRUSES
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: octapeptide antigen
 ; SEQ ID NO: 314-268-36

Query Match 56.0%; Score 51; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO: 314-268-118
 ; Sequence 118, Application US/09314268
 ; Patent No. 6346377

Query Match 57.1%; Score 52; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO: 314-268-33

Db 4 PKPSPWAP 11
 Db 158 SPWAPKHH 14

RESULT 14
 US-09-314-268-118
 ; Sequence 118, Application US/09314268
 ; Patent No. 6346377

GENERAL INFORMATION:
i APPLICANT: Doorbar, John
i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
i FILE REFERENCE: 3789/80902
i CURRENT APPLICATION NUMBER: US/09/314,268
i CURRENT FILING DATE: 1999-03-19
i EARLIER APPLICATION NUMBER: 09/314,268
i EARLIER FILING DATE: 1999-05-18
i NUMBER OF SEQ ID NOS: 179
i SOFTWARE: PatentIn Ver. 2.1
i SEQ ID NO: 118
i LENGTH: 34
i TYPE: PRT
i ORGANISM: Human papillomavirus type 35
US-09-314-268-118

Query Match Score 51; DB 4; Length 34;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 KPSPWAKPR 15
| : | : | : |
Db 1 KPAWPWPKPR 11

RESULT 15
US-09-314-268-34
Query Match Score 51; DB 4; Length 34;
Best Local Similarity 56.0%; Pred. No. 1.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 5 KPSPWAKPR 15
| : | : | : |
Db 1 KPAWPWPKPR 11

GENERAL INFORMATION:
i APPLICANT: Doorbar, John
i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
i FILE REFERENCE: 3789/80902
i CURRENT APPLICATION NUMBER: US/09/314,268
i CURRENT FILING DATE: 1999-03-19
i EARLIER APPLICATION NUMBER: 09/314,268
i EARLIER FILING DATE: 1999-05-18
i NUMBER OF SEQ ID NOS: 179
i SOFTWARE: PatentIn Ver. 2.1
i SEQ ID NO: 34
i LENGTH: 8
i TYPE: PRT
i ORGANISM: Artificial Sequence
i FEATURE:
i OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-09-314-268-34

Query Match Score 50; DB 4; Length 8;
Best Local Similarity 54.9%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5 KPSWAKP 12
| : | : | : |
Db 1 KPSWAKP 8

This Page Blank (uspto)

SUMMARIES									
Result No.	Score	Query Match Length	DB ID	Description	% Gaps				
1	91	100_0	15_14	US-10-008-524A-167	Sequence 167 , App	0	0	0	0;
2	91	100_0	25_14	US-10-008-524A-4	Sequence 4 , Appli	0	0	0	0;
3	91	100_0	26_14	US-10-350-719-4	Sequence 4 , Appli	0	0	0	0;
4	91	100_0	92_14	US-10-008-524A-90	Sequence 90 , Appli	0	0	0	0;
5	91	100_0	92_15	US-10-350-719-90	Sequence 90 , Appli	0	0	0	0;
6	82	90_1	15_15	US-10-350-719-167	Sequence 167 , App	0	0	0	0;
7	75	82_4	12_14	US-10-008-524A-168	Sequence 168 , App	0	0	0	0;
8	75	82_4	12_15	US-10-350-719-168	Sequence 168 , App	0	0	0	0;
9	68	74_7	35_14	US-10-008-524A-117	Sequence 117 , App	0	0	0	0;
10	68	74_7	35_15	US-10-350-719-117	Sequence 117 , App	0	0	0	0;
11	55	60_4	113_12	US-10-424-519-2504	Sequence 254034 ,	0	0	0	0;
12	53	58_2	55_14	US-10-029-386-31681	Sequence 31681 , A	0	0	0	0;
13	52	57_1	8_14	US-10-008-524A-31	Sequence 31 , Appli	0	0	0	0;
14	52	57_1	8_14	US-10-008-524A-33	Sequence 33 , Appli	0	0	0	0;
15	52	57_1	8_14	US-10-008-524A-17	Sequence 37 , Appli	0	0	0	0;

FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 4
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-4

Query Match 100.0%; Score 91; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.00022; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RP1PKPSPWAKKHR 15
Db 1 RP1PKPSPWAKKHR 15

RESULT 5
US-10-350-719-90
; Sequence 90, Application US/10350719
; Publication No. US20030319726A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 90
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-90

Query Match 100.0%; Score 91; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.00022; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RP1PKPSPWAKKHR 15
Db 1 RP1PKPSPWAKKHR 15

RESULT 6
US-10-350-719-167
; Sequence 167, Application US/10350719
; Publication No. US20030319726A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: GB0018140.4
; PRIOR FILING DATE: 2000-07-24
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 167
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-167

Query Match 90.1%; Score 82; DB 15; Length 15;
Best Local Similarity 93.3%; Pred. No. 0.0017; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RP1PKPSPWAKKHR 15
Db 1 RP1PKPSPWAKKHR 15

RESULT 4
US-10-008-524A-90
; Sequence 90, Application US/10008524A
; Publication No. US20030175682A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 90
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-90

```

RESULT 7
US-10-008-524A-168
Sequence 168, Application US/10008524A
Publication No. US/003017568/AA1
GENERAL INFORMATION:
APPLICANT: Doobar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
FILE REFERENCE: 18396/1074
CURRENT APPLICATION NUMBER: US/10/008,524A
PRIORITY FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: 09/314,268
PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 168
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-10-008-524A-168

RESULT 8
Query Match 82.4%; Score 75; DB 14; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.01; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 4 PKPSPWAKKHR 15
Db 1 PKPSPWAKKHR 12

RESULT 9
Query Match 82.4%; Score 75; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.01; Indels 0; Gaps 0;
Matches 12; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 4 PKPSPWAKKHR 15
Db 1 PKPSPWAKKHR 12

RESULT 10
US-10-350-719-117
Query Match 74.7%; Score 68; DB 14; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 11
US-10-350-719-117
Query Match 74.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 12
US-10-350-719-117
Query Match 74.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 13
US-10-350-719-117
Query Match 74.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 14
US-10-424-599-254004
Query Match 74.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 15
US-10-424-599-254004
Query Match 74.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.17; Indels 0; Gaps 0
Matches 11; Conservative 0; Mismatches 0; Delins 0; Gaps 0
QY 5 KPSPWAKKHR 15
Db 1 KPSPWAKKHR 11

RESULT 16
US-10-008-524A-117
Sequence 117, Application US/10008524A
Publication No. US/003017568/AA1
GENERAL INFORMATION:
APPLICANT: Doobar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
FILE REFERENCE: 18396/1074
CURRENT APPLICATION NUMBER: US/10/008,524A
PRIORITY FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: 09/314,268
PRIOR FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 117
LENGTH: 35
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-10-008-524A-117

```

OTHER INFORMATION: Clone ID: PAT_MRT3847_71461C.1.pep
US-10-0424-599-254084

Query Match Score 55; DB 12; Length 113;
Best Local Similarity 76.9%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPIPKPSWAPKK 13
Db 34 RPFPFSPSWAPRK 46

RESULT 12
US-10-029-31681
Sequence 31681, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
FILE REFERENCE: AROMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029 386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NCS: 34288
SOFTWARE: Anamax Sequence Listing Engine vers. 1.1
SEQ ID NO 31681
LENGTH: 55

TYPE: PRT
ORGANISM: Homo sapiens

FEATURE:
OTHER INFORMATION: MAP TO AC008974 6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.87
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.55
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.8
OTHER INFORMATION: SWISSPROT HIT: Q28084, EVALUE 1.20e+00
US-10-029-31681

Query Match Score 53; DB 14; Length 55;
Best Local Similarity 57.1%; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 RPIPKPSWAPKKH 14
Db 7 RPVPGASFWPGSH 20

RESULT 13
US-10-008-524A-31
Sequence 31, Application US/10008524A
Publication No. US20030175682A1
GENERAL INFORMATION:
APPLICANT: Doorbar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
FILE REFERENCE: 18396/1074
CURRENT APPLICATION NUMBER: US/10/008, 524A
PRIORITY FILING DATE: 2002-08-13
PRIOR APPLICATION NUMBER: 09/314, 268
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 31
LENGTH: 8

TYPE: PRT
ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: Octapeptide antigen

US-10-008-524A-31

Query Match Score 52; DB 14; Length 8;
Best Local Similarity 57.1%; Pred. No. 1e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PWAPKKHR 15
Db 1 PWAPKKHR 8

Search completed: May 27, 2004, 16:34:18
Job time : 31.7308 secs

This Page Blank (USPfO)

Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:13:48 ; Search time: 10.0962 Seconds
 (without alignments)
 142.913 Million cell updates/sec

Title: US-10-008-524A-167

Perfect score: 91

Sequence: 1.RP1PKPSEWAPCKHR.15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	95	1.W4WLHS	E4 protein - human
2	74	81.3	96	1.W4WLJ35	E4 protein - human
3	61	67.0	87	1.W4WL51	E4 protein - human
4	56	61.5	102	1.W4WL31	E4 protein - human
5	54	59.3	330	2.T05717	probable extensin
6	54	59.3	1553	2.T03301	rab3 effector prot
7	51	56.0	1060	2.A10201	beta-galactosidase
8	49	53.8	325	2.S57977	CCH zinc finger p
9	48	52.7	88	1.W4WL18	E4 protein - human
10	48	52.7	721	2.C84677	probable membrane
11	48	52.7	992	2.T38817	hypothetical prote
12	47	51.6	94	1.W4WL39	E4 protein - human
13	47	51.6	378	1.A40004	histidine decarbox
14	47	51.6	846	2.S52418	GTP-binding regula
15	47	51.6	925	2.T07713	probable ABC-type
16	46	50.5	230	2.S72714	LepB1170_F2_64_pro
17	46	50.5	474	2.T34193	G_protein_coupled
18	46	50.5	688	2.B64103	glycine-tRNA ligas
19	46	50.5	1007	2.E72489	hypothetical prote
20	46	50.5	1039	2.A34269	integrin alpha-2b
21	45	49.5	283	2.S13383	hydroxyproline-ric
22	45	49.5	389	2.T05782	hypothetical prote
23	45	49.5	406	2.T20330	hypothetical prote
24	44	48.4	105	1.SNSHH4	nonhistone chromos
25	44	48.4	308	2.A38582	pollen allergen PI
26	44	48.4	372	2.T01600	hypothetical prote
27	44	48.4	798	2.T50514	hypothetical prote
28	44	48.4	981	2.T16060	hypothetical prote
29	43.5	47.8	143	2.D83148	hypothetical prote

ALIGNMENTS

RESULT 1

W4WLHS

B4 protein - human papillomavirus type 16

C;Species: human papillomavirus type 16

C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 20-Aug-1999

C;Accession: A2355; T10425

R;Seedorff, K.; Kramer, G.; Durst, M.; Suhai, S.; Rowekamp, W.G.

Virology 145: 181-85, 1985

A;Title: Human papillomavirus type 16 DNA sequence.

A;Reference number: A22355; MUID:85246220; PMID:290099

A;Molecule type: DNA

A;Accession: A22355

A;Cross-references: GB:K02718; NID:9333031; PIDN:AAA46937.1; PID:g459913

R; Kennedy, I.M.; Haddow, J.K.; Clements, J.B.

J. Virol. 65, 2093-2097, 1991.

A;Title: A negative element in the human papillomavirus type 16 genome acts at the lev

A;Reference number: Z17014; MUID:91162763; PMID:1848319

A;Accession: T10425

A;Status: preliminary; translated from GB/EMBL/DDBJ

C;Keywords: early protein

Query Match 100.0%; Score 91; DB 1; Length 95;

Best Local Similarity 100.0%; Pred. No. 1.3e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1.RPIPKSPWPKKFR 15

Db 29.RPIPKSPWPKKFR 43

RESULT 2

W4WL35

E4 protein - human papillomavirus type 35

C;Species: human papillomavirus type 35

C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 27-Jan-1995

C;Accession: C40824

R;Marich, J.E.; Portslsler, A.V.; Rice, S.M.; McGraw, K.A.; Dubensky, T.W.

Virology 186, 770-776, 1992

A;Title: The phylogenetic relationship and complete nucleotide sequence of human papill

A;Reference number: A40824; MUID:92124753; PMID:1310198

A;Accession: C40824

A;Status: translation not shown

A;Molecule type: DNA

A;Accession: 1-96 <M>

A;Cross-references: GB:W74117
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match Score 74; DB 1; Length 96;
 Best Local Similarity 80.0%; Pred. No. 0.0026;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPIPKSPWAPKHK 15
 | | | : | | | : | |
 Db 29 RPIPKPAPWAPQPKR 43

RESULT 3
 W4WL51
 E4 protein - human papillomavirus type 51.
 C;Species: human papillomavirus type 51
 C;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1999 #sequence_revision 31-Mar-1999 #text_change 27-Jan-1995
 C;Accession: C40415
 R:Lungu, O.; Crum, C.P.; Silverstein, S.J.,
 J.Virol. 65, 4216-4225, 1991
 A;Title: Biologic properties and nucleotide sequence analysis of human papillomavirus type 51
 A;Reference number: A40415; PMID:91303675; PMID:169326
 A;Accession: C40415
 A;Status: translation not shown
 B;Molecule type: DNA
 A;Residues: 1-87 <LNU>
 A;Cross-references: GB:W62877
 C;Superfamily: Papillomavirus E4 protein
 C;Keywords: early protein

Query Match Score 61; DB 1; Length 87;
 Best Local Similarity 73.3%; Pred. No. 0.13%;
 Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPIPKSPWAPKHK 15
 | | | : | | | : | |
 Db 24 RPIPLPPAWAPKPR 38

RESULT 4
 W4WL31
 E4 protein - human papillomavirus type 31
 C;Species: human papillomavirus type 31
 C;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1999 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
 C;Accession: E32444
 R:Goldsbrough, M.D.; DiSilvestre, D.; Temple, G.F.; Lorincz, A.T.,
 Virology 171, 306-311, 1989
 A;Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associating virus
 A;Reference number: A94398; PMID:89239478; PMID:2545036
 A;Accession: E32444
 A;Status: translation not shown
 B;Molecule type: DNA
 A;Residues: 1-102 <GOI>
 A;Cross-references: GB:J04353; NID:G333048; PIDN:AAA46949.1; PID:9459915
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match Score 56; DB 1; Length 102;
 Best Local Similarity 81.8%; Pred. No. 0.71%;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 IPKPSWAPKK 13
 | | | : | | |
 Db 32 IPKPAPWAPVK 42

RESULT 5
 T05717
 probable extensin - barley (fragment)

C;Species: Hordeum vulgare (barley)
 C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 20-Jun-2000
 C;Accession: T05717
 R;Doan, D.N.P.; Sturaro, M.; Olsen, O.A.
 submitted to the EMBL Data Library, July 1997
 A;Description: Characterization of a mucellar cDNA encoding a probable extensin from de
 A;Reference number: Z15429
 A;Accession: T05717
 A;Status: Preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-330 <DOA>
 A;Cross-references: EMBL:Z98204; PIDN:CAB10894.1
 C;Genetics:
 C;Gene: ex1
 A;Map Position: 2
 C;Superfamily: hydroxyproline-rich glycoprotein

Query Match Score 54; DB 2; Length 330;
 Best Local Similarity 75.0%; Pred. No. 4.1%;
 Matches 9; Conservative 2; Wismatches 1; Indels 0; Gaps 0;

Qy 1 RPIPKSPWAPK 12
 | : | | | | |
 Db 200 KPKPSSPPAKP 211

RESULT 6
 T03301
 rab3 effector protein Rim - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
 C;Accession: T03301
 R:Wang, Y.; Okamoto, M.; Schmitz, F.; Hofmann, K.; Sudhof, T.C.,
 Nature 388, 593-598, 1997
 A;Title: Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion
 A;Reference number: Z14897; PMID:97394473; PMID:952191
 A;Accession: T03301
 A;Status: Preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-153 <WAN>
 A;Cross-references: EMBL:AF007836; NID:G2317777; PIDN:AB66703.1; PID:92317778
 A;Experimental source: tissue-type brain
 C;Genetics:
 C;Function:
 A;Description: Rim protein is proposed as Rab3-dependent regulator of synaptic-vesicle
 C;Keywords: GTP binding; zinc finger

Query Match Score 54; DB 2; Length 1553;
 Best Local Similarity 53.8%; Pred. No. 18%;
 Matches 7; Conservative 5; Wismatches 1; Indels 0; Gaps 0;

Qy 2 PIPKPSWAPKKH 14
 | : | | | : | : |
 Db 890 PLPQPSPEMPRKH 902

RESULT 7
 AI0201
 beta-D-galactosidase (EC 3.2.1.23) [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C;Accession: AI0201
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dougan, G.;
 ill, M.; Rutherford, K.; Simmonds, M.; Shelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague
 A;Reference number: AB0001; MUID:21470413; PMID:11586360
 A;Accession: AI0201
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues: 1-1060 <KDR>

A;Cross-references: GB:AL590842; PIDN:CARC90476.1; PID:gi15979691; GSPDB:GN00175 C;Gene: lacZ C;Superfamily: beta-galactosidase C;Keywords: glycosidase; hydrolase	Best Local Similarity 63.6%; Pred. No. 74; Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Query Match 56.0%; Score 51; DB 2; Length 1060; Best Local Similarity 58.3%; Pred. No. 31; Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0; RESULT 10 C84677	Probable membrane transporter [imported] - <i>Arabidopsis thaliana</i> C;Species: <i>Arabidopsis thaliana</i> (mouse ear cress) C;Accession: C84677 C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
QY 4 PKPSPWAPKXR 15 Db 723 PRTPWSPQR 734	R;Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, L.; Tallon, L.; euss, D.; Niezman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, Nature 402, 761-765, 1999 A;Title: Sequence and analysis of chromosome 2 of the plant <i>Arabidopsis thaliana</i> . A;Reference number: A84420; PMID:2008487; PMID:1061197 A;Status: preliminary A;Accession: C84677 A;Residues: 1-721 <STO> A;Cross-references: GB:AE002093; NID:93860251; PIDN: AAC73019.1; GSPDB:GN00139 C;Genetics: A;Gene: At2g2810 A;Map position: 2
RESULT 8 S57977 CCCH zinc finger protein CTTH1 - yeast (<i>Saccharomyces cerevisiae</i>) N;Alternate names: protein YD8358.07c; protein YDR151c C;Accession: S57977; JCS501 C;Title: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999 R;Murphy, L.; Richards, C.; Harris, D. submitted to the EMBL Data Library, July 1995 A;Reference number: S57977 A;Molecule type: DNA A;Accession: S57977 A;Residues: 1-325 <MUR> A;Cross-references: EMBL:250046; NID:g899393; PIDN:CAA90373.1; PID:9899400; MIPS:YDR151c A;Experimental source: strain AB972 R;Thompson, M.J.; Lai, W.S.; Taylor, G.A.; Blackshear, P.J. Gene 174, 225-233, 1996 A;Title: Cloning and characterization of two yeast genes encoding members of the CCCH cl A;Reference number: JC5001; PMID:97045817; PMID:8890739 A;Molecule type: DNA A;Residues: 1-141; RV' 144-325 <THO> A;Cross-references: GB:i42133; NID:gi1020082; PIDN:AB39897.1; PID:gi1020083 C;Comment: This protein belongs to the CCCH-type zinc finger protein family, and is a su C;Genetics: A;Gene: SGD:CTTH1 A;Cross-references: SGD:S0002558; MIPS:YDR151c A;Map position: 4R C;Keywords: zinc finger F;322-136/Region: nuclear location signal	Query Match 52.7%; Score 48; DB 2; Length 721; Best Local Similarity 63.6%; Pred. No. 55; Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0; C84677 QY 1 RPPIPKSPWP 11 Db 7 KPGPRPGPWP 17
QY 4 PKPSPWAPKXR 13 Db 79 PRPSFWLPSK 88	T38817 hypothetical protein SPAC4F10.13C - fission yeast (<i>Schizosaccharomyces pombe</i>) C;Species: <i>Schizosaccharomyces pombe</i> C;Accession: T38817 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 R;Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, September 1997 A;Reference number: Z21813 A;Accession: T38817 A;Status: preliminary; translated from GB/EMBL/DBJ A;Molecule type: DNA A;Residues: 1-992 <CONS> A;Experimental source: strain 972h-; cosmid c4F10 C;Genetics: A;Gene: SPDB:SPAC4F10.13C A;Map position: 1 A;Map: 13/2
RESULT 9 W4H118 B4 protein - human papillomavirus type 18 C;Species: human papillomavirus type 18 C;Accession: E26251 R;Cole, S.T.; Danos, O. J. Mol. Biol. 193, 599-608, 1987 A;Title: Nucleotide sequence and comparative analysis of the human papillomavirus type 1 A;Reference number: A92937; PMID:3039146 A;Molecule type: DNA A;Residues: 1-88 <COL> A;Cross-references: GB:X05015; NID:960975; PIDN:CAA28668.1; PID:gi60980 C;Superfamily: Papillomavirus E4 protein C;Keywords: early protein	Query Match 52.7%; Score 48; DB 2; Length 992; Best Local Similarity 69.2%; Pred. No. 74; Matches 9; Conservative 0; Mismatches 2; Indels 2; Gaps 1; C;Accession: E26251 QY 4 PKDSPW- APKXH 14 Db 622 PKESPWKLPPKH 634
Query Match 52.7%; Score 48; DB 1; Length 88;	RESULT 12 W4H118 B4 protein - human papillomavirus type 39 C;Species: human papillomavirus type 39 C;Note: host Homo sapiens (man) C;Accession: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 27-Jan-1995

C;Accession: E38502
 R;Volpers, C.; Streeck, R.E.
 Viralp. 191, 419-423, 1991
 A;Title: Genome organization and nucleotide sequence of human papillomavirus type 39.
 A;Reference number: A38502; MUID:91135017; PMID:1847266
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-94 <VO1>
 A;Cross-references: EMBL:M38155
 C;Keywords: Papillomavirus B4 protein
 C;Keywords: early protein

Query Match 51.6%; Score 47; DB 1; Length 94;
 Best Local Similarity 69.2%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 4;
 Indels 0; Gaps 0;

Qy 1 RPPIPKSPWAPKK 13
 Db 29 RPPIPPQPHAPKK 41

RESULT 13

A40004 histidine decarboxylase (EC 4.1.1.22) - Enterobacter aerogenes
 C;Species: Enterobacter aerogenes
 C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 18-Jun-1999
 C;Accession: A40004
 R;Kamath, A.V.; Vaaler, G.L.; Snell, E.E.
 J. Biol. Chem. 266, 9432-9437, 1991
 A;Title: Pyridoxal phosphate-dependent histidine decarboxylases. Cloning, sequencing, and enzymes.
 B;Reference number: A40004; MUID:91236707; PMID:2033044
 A;Accession: A40004
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-378 <KM>
 A;Cross-references: GB:MG62745; NID:943553; PIDN:AAA24802.1; PID:9435594
 C;Superfamily: Klebsiella histidine decarboxylase
 C;Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate
 F;233/Binding site: Pyridoxal Phosphate (Lys) (covalent) #status predicted

Query Match 51.6%; Score 47; DB 1; Length 378;
 Best Local Similarity 72.7%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 3;
 Indels 0; Gaps 0;

Qy 4 PKPSPWAPKKH 14
 Db 329 PKPSEWWKKH 339

RESULT 14

S52418 GTP-binding regulatory protein Gs alpha-XL chain - rat
 N;Alternative names: G protein XL-alpha-s
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 14-Jul-1995 #sequence_revision 10-Nov-1995 #text_change 02-Feb-2001
 C;Accession: S52418
 R;Kehlenbach, R.H.; Matthey, J.; Huttner, W.B.
 Nature 372, 804-809, 1994
 A;Title: XL-alpha-s is a new type of G protein.
 B;Reference number: S52418; MUID:95089824; PMID:7997272
 A;Accession: S52418
 B;Molecule type: mRNA
 B;Residues: 1-846 <KEF>
 A;Cross-references: EMBL:X84047; NID:9642267; PIDN:CAA58866.1; PID:g642268

Nature 375, 253, 1995

A;Title: Correction: XLalphas is a new type of G protein.

A;Contents: annotation; assignment of start_codon

A;Note: experimental data from this paper suggest that the translation is initiated at p
 C;Keywords: GTP binding; nucleotide binding; p-loop; signal transduction

Scoring table:	BLOSUM62					
Gapop:	Gapext 0.5					
Searched:	141681 seqs, 52070155 residues					
Total number of hits satisfying chosen parameters:	141681					
Minimum DB seq length:	0					
Maximum DB seq length:	2000000000					
Post-processing:	Minimum Match 0%					
	Maximum Match 100%					
Database :	SwissProt_42:*					
Pred.	No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.					
SUMMARIES						
Result No.	Query	Score	Match	Length	DB ID	Description
1	91	100.0	95	1	VE4 HPV16	P06222 human papill
2	74	81.3	96	1	VE4 HPV35	P27224 human Papil
3	61	67.0	87	1	VE4 HPV51	P26548 human Papil
4	56	61.5	102	1	VE4 HPV31	P17384 human Papil
5	54	59.3	1615	1	RIM1 RAT	Q9jir4 rattus norv
6	54	59.3	1692	1	RIM1 HUMAN	Q86x5 homo sapien
7	49	53.8	325	1	CTH1 YEAST	P4736 saccharomye
8	48	53.3	1513	1	GRLF HUMAN	Q9nry4 homo sapien
9	48	52.7	88	1	VE4 HPV18	P06191 human papil
10	48	52.7	1532	1	RIM2 MOUSE	Q9eqz7 mus musculu
11	47	51.6	94	1	VE4 HPV39	P24831 human papil
12	47	51.6	377	1	DCHS_ENTAE	P28577 enterobacte
13	46	50.5	454	1	CT81 HUMAN	Q9hiq7 homo sapien
14	46	50.5	688	1	SYGB_HAEM	P42822 haemophilus
15	46	50.5	689	1	SYGB_PASMU	P51905 pastorella
16	46	50.5	755	1	MTS1_HUMAN	Q4312 homo sapien
17	46	50.5	755	1	MTS1_MOUSE	Q8r154 mus musculu
18	46	50.5	1039	1	ITAB_HUMAN	P0514 homo sapien
19	45	49.5	153	1	PE1_ANOGA	P6217 anophelis g
20	45	49.5	219	1	MGH1_HUMAN	Q9i213 homo sapien
21	45	49.5	283	1	EXTIN_SORBI	P24152 sorghum bic
22	45	49.5	361	1	IHA_TRIVU	07755 trichosurus
23	45	49.5	461	1	Y514_HUMAN	06129 homo sapien
24	45	49.5	493	1	ANL2_HUMAN	Q9uk9 homo sapien
25	45	49.5	1188	1	RIM2_HUMAN	P1126 homo sapien
26	45	49.5	1555	1	RIM2 RAT	Q9jiz1 rattus norv
27	44	48.4	104	1	HG15_CHICK	P12902 gallus galli
28	44	48.4	174	1	R122_SECCSE	Q00536 secale cere
29	44	48.4	308	1	MP5A_LOLPR	Q41240 lolium per
30	43	48.4	670	1	ZN16_HUMAN	P17020 homo sapien
31	43	47.8	814	1	TOP1_SCHPO	P0799 schizosacch
32	43	47.3	111	1	UL91_HCMV	P16797 human cyt
33	43	47.3	128	1	VDBP_CERV	P05398 carnation e

DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable E4 protein.
 RT Human papillomavirus type 35.
 OS Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus
 NCBI_TaxID=10587;
 CX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92124753; PubMed=1310198;
 RA Marich J.E., Pontsler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
 RT "The phylogenetic relationship and complete nucleotide sequence of
 human papillomavirus type 35.",
 RL Virology 186:770-776 (1992).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL; M74117; AAA46965.2; -.
 DR PIR; C40B824; WAWL35.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 DR Early protein.
 SQ SEQUENCE 96 AA; 10597 MW; AE4524418CD26F7C CRC64;
 Query Match Score 74; DB 1; Length 96;
 Best Local Similarity 80.0%; Pred. No. 0.0021;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RP1PKPSPWAPKKHR 15
 Db 29 RP1PKPAPWAPQKPR 43

Query Match Score 74; DB 1; Length 96;
 Best Local Similarity 80.0%; Pred. No. 0.0021;
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 RP1PKPSPWAPKXHR 15
 Db 29 RP1PKPAPWAPQKPR 43

RESULT 3
 VE4 HPV51 STANDARD PRT; 87 AA.
 AC P26548;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-AUG-1992 (Rel. 23, Last annotation update)
 DE Probable E4 protein.
 OS Human papillomavirus type 51.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 NCBI_TaxID=10595;
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE=9103675; PubMed=1649326;
 RT "Biologic properties and nucleotide sequence analysis of human
 papillomavirus type 51.";
 RT J. Virol. 65:4216-4225(1991).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC DR NOT_ANNOTATED_CDS.
 DR PIR; CA0415; W4W151.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 DR Early protein.
 SQ SEQUENCE 87 AA; 9941 MW; 5F3DC38FF86BF3990 CRC64;

Query Match Score 61; DB 1; Length 87;
 Best Local Similarity 73.3%; Pred. No. 0.094; 4; Mismatches 0; Indels 0; Gaps 0;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 RP1PKPSPWAPKXHR 15
 Db 24 RP1PLPPWAPFKPR 38

RESULT 4
 VE4 HPV31 STANDARD PRT; 102 AA.
 AC P17394;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE Probable E4 protein.
 OS Human papillomavirus type 31.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 NCBI_TaxID=10585;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=81299478; PubMed=2545036;
 RA Goldsbrough M.D., Disi'vestre D., Temple G.F., Lorincz A.T.;
 RT "Nucleotide sequence of human papillomavirus type 31: a cervical
 neoplasia-associated virus.",
 RL Virology 171:306-311 (1989).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC DR EMBL; J0453; AAA46949.1; -.
 DR PIR; E32444; WAWL31.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 DR Early protein.
 SQ SEQUENCE 102 AA; 11284 MW; 04E3C9EBABC5CA6C CRC64;
 Query Match Score 56; DB 1; Length 102;
 Best Local Similarity 81.8%; Pred. No. 0.49; 1; Mismatches 1; Indels 0; Gaps 0;
 Matches 9; Conservative 81; Mismatches 1; Indels 0; Gaps 0;
 Qy 3 IPKPKSPWAPKK 13
 Db 32 IPKPKAPWAPVK 42

RESULT 5
 RIM1 RAT STANDARD PRT; 1615 AA.
 ID RIM1_RAT
 AC Q9JX4; Q35168;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Regulating synaptic membrane exocytosis protein 1 (Rab3-interacting
 molecule 1) (RIM 1).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Sciurognathi;
 OC Mammalia; Eutheria; Rodentia; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN RP SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RAB3A AND RAB3C.
 RC TISSUE=Brain;
 RX MEDLINE=97394473; PubMed=9252191;
 RA Wang Y., Okamoto M., Schmitz F., Hoffmann K., Suedhof T.C.;

Comment=Additional isoforms seem to exist;

CC IsoId=Q86UR5-1; Sequence=Displayed;

CC Name=2; Synonyms=RIM short form;

CC IsoId=Q86UR5-2; Sequence=VSP_008160; VSP_008165;

CC Name=3; Synonyms=RIM long form; Rab3 interacting protein variant

2; IsoId=Q86UR5-3; Sequence=VSP_008160; VSP_008164; VSP_008167;

CC Name=4; Synonyms=Rab3 interacting protein variant 1;

CC IsoId=Q86UR5-4; Sequence=VSP_008166; VSP_008167;

CC Name=5; Synonyms=Rab3 interacting protein variant 3;

CC IsoId=Q86UR5-5; Sequence=VSP_008161; VSP_008163; VSP_008170;

CC Name=6; Synonyms=Rab3 interacting protein variant 4;

CC IsoId=Q86UR5-6; Sequence=VSP_008164; VSP_008168; VSP_008169;

CC Name=7; Synonyms=Rab3 interacting protein variant 5;

CC IsoId=Q86UR5-7; Sequence=VSP_008161; VSP_008164; VSP_008167;

CC Name=8; Synonyms=Rab3 interacting protein variant 6;

CC IsoId=Q86UR5-8; Sequence=VSP_008161; VSP_008162; VSP_008169;

CC -!- TISSUE_SPECIFICITY: Detected in brain and retina.

CC -!- DISEASE: Defects in RIMS1 are a cause of autosomal dominant cone-rod dystrophy (CORD7) [MIM:603649]. CORD7 is characterized by early loss of visual acuity and color vision, followed by night blindness and peripheral visual field loss. The onset of reduced color vision and visual acuity varies between the ages of 20 and 40 years.

CC -!- SIMILARITY: Contains 2 C2 domains.

CC -!- SIMILARITY: Contains 1 FYVE-type zinc finger.

CC -!- SIMILARITY: Contains 1 PDZ/DIR domain.

CC -!- SIMILARITY: Contains 1 Rab-binding (RabBD) domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AY190519; AAO38848.1; -.

CC DR EMBL; AB045726; BAB87121.1; -.

CC DR EMBL; AB051866; BAB87242.1; -.

CC DR EMBL; AB0023318; BAA20738.1; ALT_INIT.

CC DR EMBL; AF263305; AAG231162.1; -.

CC DR EMBL; AF263306; AAG231163.1; -.

CC DR EMBL; AF263307; AAG231164.1; -.

CC DR EMBL; AF263308; AAG231155.1; -.

CC DR EMBL; AF263309; AAG231166.1; -.

CC DR EMBL; AF263310; AAG231167.1; -.

CC DR HSSP; P21707; 1RSY.

CC DR Genew; HGNC:17282; RIMS1.

CC DR MIM; 606629; -.

CC DR MIM; 603649; -.

CC DR InterPro; IPR000008; C2.

CC DR InterPro; IPR008973; C2_CalB.

CC DR InterPro; IPR001478; PDZ.

CC DR InterPro; IPR033315; RPH3A_effector.

CC DR InterPro; IPR003036; Znf_FYVE.

CC DR Pfam; PF00168; C2; 2.

CC DR Pfam; PF00595; PDZ; 1.

CC DR Pfam; PF02318; RPH3A_effector; 1.

CC DR SMART; SM00233; C2; 2.

CC DR PROSITE; PS50104; C2_DOMAIN_2; 1.

CC DR PROSITE; PS50106; PDZ; 1.

CC DR PROSITE; PS50116; RABBD; 1.

CC DR DOMAIN; PS50116; ZF_FYVE; 1.

CC KW Metal-binding Zinc_Repeat; zinc-finger; zinc-finger; Alternative splicing;

FT DOMAIN; Disease_mutation.

FT DOMAIN; 22 182

FT ZN_FING 110 170

FT DOMAIN; 605 691

RAB-BINDING.

FYVE-TYPE.

PDZ.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

Query Match Score 49; DB 1; Length 325;
 Best Local Similarity 70.0%; Pred. No. 12;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy	4 PKPSPWPKK 13
Db	79 PRPSPWLP SK 88

RESULT 8
GRLE_HUMAN STANDARD; PRT; 1513 AA.
 ID Q9NRY4; Q14452; Q9C0E1;
 AC 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Glucocorticoid receptor DNA binding factor 1 (Glucocorticoid receptor
 GRF1) OR GRF1 OR KIAA1722.
 DN Homo sapiens (Human).
 OS Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

[1] RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Breast cancer;
 MEDLINE=2101021; PubMed=11054565;
 RA Tikro A., Czeckay S., Vlars C., White S., Heath J.K., Arden K.,
 RA Maruta H.;
 RT "p190-A, a human tumor suppressor gene, maps to the chromosomal region
 19q13.3, that is reportedly deleted in some gliomas.";
 RT Gene 257:23-31(2000).
 RN [2] RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 MEDLINE=2108232; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.,
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro";
 RT DNA Res. 7:347-355(2000).
 RN [3] RP SEQUENCE OF 377-1453 FROM N.A. (ISOFORM 1), AND FUNCTION.
 RC TISSUE=Breast cancer;
 MEDLINE=9137352; PubMed=1894621;
 RA LeClerc S., Palaniswami R., Xie B.X., Govindan M.V.,
 RT "Molecular cloning and characterization of a factor that binds the
 human glucocorticoid gene and represses its expression.";
 RL J. Biol. Chem. 266:17333-17340(1991).
 RT :- FUNCTION: Represses transcription of the glucocorticoid receptor

CC by binding to the cis-acting regulatory sequence 5' -
 CC GAGAAAAAGAAACTGGAAATTC-3 . May participate in the regulation of
 CC retinal development and degeneration. May transduce signals from
 CC p21-ras to the nucleus, acting via the ras GTP-ase activating
 protein (GAP). May also act as a tumor suppressor.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and nuclear (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=2;
 CC Name=;
 CC IsoId=Q9NRY4-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q9NRY4-2; Sequence=VSP 007105;
 CC Note=No experimental confirmation available;
 CC -!- PTM: Tyrosine phosphorylated (By similarity).
 CC -!- SIMILARITY: Contains 1 Rho-GAP domain.
 CC -!- SIMILARITY: Contains 4 FF domains.
 CC -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
 CC in positions 533: 540; 607 and 614.
 CC -!- CAUTION: Ref.3 sequence differs from that shown due to frameshifts
 CC in positions 389, 533, 540; 607, 614; 1167, 1241, 1292, 1334 and
 CC 1446.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licensee@isb-sib.ch).
 CC
 DR EMBL; AF159851; AAP03861; ALT_FRAME.
 DR EMBL; AB05150; BA21813; 2; -.
 DR EMBL; MT2077; AAA5618; 1; ALT_FRAME.
 DR TRANSFAC; T00940;
 DR Genew; HGNC:4591; GRLF1.
 DR MIM: 60277; -.
 DR GO:0005634; C:nucleus; IC.
 DR GO:0016564; F:transcriptional repressor activity; IDA.
 DR GO:0016481; F:negative regulation of transcription; IDA.
 DR GO:0000074; P:regulation of cell cycle; NAS.
 DR InterPro; IPR002713; FF.
 DR InterPro; IPR001806; Ras_transfmg.
 DR InterPro; IPR008936; RhogAP.
 DR InterPro; IPR00198; RhogAP.
 DR Pfam; PF01846; FF; 2.
 DR Pfam; PF00620; RhogAP; 1.
 DR PRINTS; PR000419; RASTRNSFRNG.
 DR SMART; SM00441; FF; 4.
 DR PROSITE; PS50338; RHOGAP; 1.
 DR GTBase activation; DNA-binding; Repressor; Transcription regulation;
 KW Anti-oncogene; Nuclear Protein; Phosphorylation; Repeat;
 KW Alternative splicing.
 FT DOMAIN 270 327 FF 1.
 FT DOMAIN 369 422 FF 2.
 FT DOMAIN 429 483 FF 3.
 FT DOMAIN 485 539 FF 4.
 FT DOMAIN 1249 1436 RHO_GAP.
 FT DOMAIN 1440 1507 PRO_RICH.
 FT VARPLIC 1492 1513 isoform 2.
 FT /FTId=VSP 007105.
 FT CONFLICT 251 251 R -> P (IN REF. 1 AND 3).
 FT CONFLICT 309 309 V -> D (IN REF. 1 AND 3).
 FT CONFLICT 362 362 S -> G (IN REF. 1 AND 3).
 FT CONFLICT 388 388 W -> M (IN REF. 3).
 FT CONFLICT 414 414 Q -> A (IN REF. 1 AND 3).
 FT CONFLICT 474 474 M -> T (IN REF. 1 AND 3).
 FT CONFLICT 978 978 C -> S (IN REF. 1).
 FT CONFLICT 1292 1292 M -> I (IN REF. 1).
 FT CONFLICT 1452 1453 PS -> RN (IN REF. 3).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AB021131; BAB18975_1; -.

EMBL: AK032619; BAC27953_1; -.

EMBL: AK083172; BAC3B794_1; -.

HSSP: P21707; 1BNW.

MGD; MGI:21552972; Rims2.

GO; GO:0005515; F:protein binding; IPI.

GO; GO:0017156; P:calcium ion dependent exocytosis; IDA.

GO; GO:0019933; P:CaM-mediated signaling; IDA.

GO; GO:0030073; P:insulin secretion; IDA.

InterPro: IPR000008; C2.

InterPro: IPR008973; C2_CalB.

InterPro: IPR01478; PDZ.

InterPro: IPR000306; Znf_FYVE.

Pfam: PF00168; C2; 2.

Pfam: PF00595; PDZ; 1.

SMART: SM00239; C2; 2.

SMART: SM00228; PDZ; 1.

PROSITE: PS50004; C2_DOMAIN_2; 2.

PROSITE: PS50106; PDZ; 1.

PROSITE: PS50916; RABD; 1.

PROSITE: PS50178; ZF_FYVE; 1.

Metal binding; Zinc; Repeat; Zinc-finger; Alternative splicing.

RP SEQUENCE FROM N.A. / ATCC 51907;
 RC KW20 / KW20 /
 RX MEDLINE=93350630; PubMed=7542900;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Fierlaiava A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Sherry R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd.";
 RL Science 269:496-512(1995).
 CC -!- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) = AMP + diphosphate
 CC + glyceryl-tRNA (Gly).
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
 CC ---
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC ---
 DR EMBL; U32774; AAC2584.1; -.
 DR PIR; B64103; B64103.
 DR HMPAP; MF_00255; -.
 DR InterPro; IPR002311; tRNA synt 2f.
 DR InterPro; IPR006194; tRNA synt_Gly.
 DR PF0209; tRNA synt_2f; 1.
 DR TIGRFAMS; TIGR00211; GLYS; 1.
 DR PROSITE; PS50861; AA_TRNA_LIGASE_II_GLYAB; 1.
 DR PR01045; tRNA synt.
 DR TIGRFAMS; TIGR00211; GLYS; 1.
 DR PROSITE; PS50861; AA_TRNA_LIGASE_II_GLYAB; 1.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 KW Complete proteome;
 SQ 688 AA; 75612 MW; 67F9ACEF0A9184 CRC64;

Query Match 50.5%; Score 46; DB 1; Length 689;
 Best Local Similarity 72.7%; Pred. No. 59;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PIPKSPWAK 12
 Db 143 PIPKPMRWAK 153

RESULT 15
 SYCB_PASMU STANDARD; PRT; 689 AA.
 ID SYGB_PASMU
 AC P57805;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Glyceryl-tRNA synthetase beta chain (EC 6.1.1.14) (Glycine-tRNA ligase
 beta chain) (GLYRS).
 GN GLYS OR PM102.
 OS Pasteurella multocida.
 OC Pasteurellaceae; Pasteurellales;
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Pasteurella.
 OX NCBI_TAXID=747;
 RN SEQUENCE FROM N.A.
 RP STRAIN=Pm70;
 RC MEDLINE=2145866; PubMed=11248100;
 RA MAY B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:12:58 ; Search time 28.8462 Seconds

(without alignments)
 164.069 Million cell updates/sec

Title: US-10-008-524A-167

Perfect score: 91

Sequence: 1 RPIPKPSPWAPKHKR 15

Scoring table: BLOSUM62

Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL 25; *

1: sp_archea;*

2: sp_bacteria;*

3: sp_fungi;*

4: sp_human;*

5: sp_invertebrate;*

6: sp_mammal;*

7: sp_mhc;*

8: sp_organelle;*

9: sp_organism;*

10: sp_plant;*

11: sp_rabbit;*

12: sp_virus;*

13: sp_vertebrate;*

14: sp_unclassified;*

15: sp_virus;*

16: sp_bacterioplasmid;*

17: sp_archaea;*

Q91tM2	tupaia	harp
Q8ZFP0	yersinia	pe
Q9gk54	macaca	fasc
Q91r55	human	papill
Q9sm80	oryza	sativ
Q8xv59	railstonia	s
Q84rp1	oryza	sativ
Q86mt6	plasmoidium	
Q8wvz3	homo	sapien
Q882z1	pseudomonas	
Q9veP4	bacteriophila	
Q7ym9	arabidopsis	
Q81p19	arabidopsis	
Q9zyY2	schizosacch	
Q36025	schizosacch	
Q7ztn2	xenopus	lae
Q9gsd5	ancylostoma	
Q9d956	mus	musculu
Q94ai8	oryza	sativ
Q63803	rattus	norv
Q9st9	arabidopsis	
Q91tp2	tupaia	harp
Q90727	human	papil
Q82001	human	papil
Q857x2	mycobacteri	
Q15150	homo	sapien
Q49630	mycobacteri	

ALIGNMENTS

RESULT 1

Q918t2	PRELIMINARY;	PRT;	95 AA.
ID	Q918t2;		
AC	Q918t2;		
DT	01-DBC-2001 (TREMBLrel.	19, Created)	
DT	01-DBC-2001 (TREMBLrel.	19, Last sequence update)	
DT	01-MAR-2002 (TREMBLrel.	20, Last annotation update)	
DE	E4	Protein (Fragment).	
GN	E4.		
OS	Human papillomavirus type 16		
OS	Viruses; dsDNA viruses; no RNA stage; Papillomaviridae;		
OC	Papillomavirus		
OC	Papillomavirus		
NCBI_TAXID	10581;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN=HPV16ECC7;		
RC	Watts K.J., Thompson C.H., Cossart Y.E., Rose B.R.;		
RA	"Sequence variation and physical state of human Papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia.";		
RT	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
RL	EMBL; AF07220; AU01405.1; -.		
DR	InterPro; IPR003861; Papilloma_E4.		
DR	Pfam; PF02711; Pap_E4; 1..		
FT	NON_TER 1_1	MN	AED17903867307CE CRC64;
SEQUENCE	95 AA;	10542	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
2	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
3	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
4	91	Q8B5P2	Q8B5P2	12	Q8B5P2	Q8B5P2	human papil
5	91	Q8B5N7	Q8B5N7	12	Q8B5N7	Q8B5N7	human papil
6	91	Q80MM0	Q80MM0	12	Q80MM0	Q80MM0	human papil
7	68	Q93z6	Q93z6	12	Q93z6	Q93z6	human papil
8	62	Q9AY37	Q9AY37	10	Q9AY37	Q9AY37	oryza sativ
9	62	Q7XC25	Q7XC25	10	Q7XC25	Q7XC25	oryza sativ
10	56	Q9PZN8	Q9PZN8	12	Q9PZN8	Q9PZN8	eastern equ
11	56	Q9PZM7	Q9PZM7	12	Q9PZM7	Q9PZM7	eastern equ
12	56	Q9PZM6	Q9PZM6	12	Q9PZM6	Q9PZM6	eastern equ
13	54	Q9AY37	Q9AY37	10	Q9AY37	Q9AY37	oryza sativ
14	54	Q9LIP1	Q9LIP1	10	Q9LIP1	Q9LIP1	oryza sativ
15	53	Q85W93	Q85W93	12	Q85W93	Q85W93	plasmoidis
16	51	Q8NEF7	Q8NEF7	4	Q8NEF7	Q8NEF7	homo sapien

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

STUNNARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
2	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
3	91	Q918t2	Q918t2	12	Q918t2	Q918t2	human papil
4	91	Q8B5P2	Q8B5P2	12	Q8B5P2	Q8B5P2	human papil
5	91	Q8B5N7	Q8B5N7	12	Q8B5N7	Q8B5N7	human papil
6	91	Q80MM0	Q80MM0	12	Q80MM0	Q80MM0	human papil
7	68	Q93z6	Q93z6	12	Q93z6	Q93z6	human papil
8	62	Q9AY37	Q9AY37	10	Q9AY37	Q9AY37	oryza sativ
9	62	Q7XC25	Q7XC25	10	Q7XC25	Q7XC25	oryza sativ
10	56	Q9PZN8	Q9PZN8	12	Q9PZN8	Q9PZN8	eastern equ
11	56	Q9PZM7	Q9PZM7	12	Q9PZM7	Q9PZM7	eastern equ
12	56	Q9PZM6	Q9PZM6	12	Q9PZM6	Q9PZM6	eastern equ
13	54	Q9AY37	Q9AY37	10	Q9AY37	Q9AY37	oryza sativ
14	54	Q9LIP1	Q9LIP1	10	Q9LIP1	Q9LIP1	oryza sativ
15	53	Q85W93	Q85W93	12	Q85W93	Q85W93	plasmoidis
16	51	Q8NEF7	Q8NEF7	4	Q8NEF7	Q8NEF7	homo sapien

RESULT 2	1 RPTPKSPNPAPKHKR 15		
QY			Gaps
Db	29 RPTPKSPWPAPKHKR 43		
QY			Gaps
Db	Q918U0	PRELIMINARY;	PRT;
AC	Q918U0;		95 AA.

OC Papillomavirus.	RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E., Craven J., Khaikin H., Feldblum T.V., Quackenbush J., White O., Salzberg S.L., SEQUENCE FROM N.A.
RP STRAIN=Asian-American variant;	RA
RC Terri M., Burk R.D.;	RA
RA "Human papillomavirus type 16 Asian-American variant." ⁵ ;	RA
RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.	RT "Oryza sativa chromosome 10 BAC OSJNBA0027P10 genomic sequence." ⁶
RL DR AC084763; AAG60194; 1.	RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR Graeme; Q9AY37;	DR
DR GO; GO:0016305; F:kinase activity, TEA.	DR
DR InterPro; IPR003409; MORN.	DR
PFam; PF02493; MORN; 7.	DR
FT SMART; SM00690; MORN; 7.	DR
KW Kinase.	KW
SQ SEQUENCE 95 AA; 10542 MW; AED17903867307CE CRC64;	SQ 467 AA; 50906 MW; 8FB85BCC9B980D73 CRC64;
Query Match Score 91; DB 12; Length 95;	Query Match Score 62; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;	Best Local Similarity 68.1%; Pred. No. 0.47;
Matches 15; Conservative 0; Mismatches 0;	Matches 10; Conservative 2; Mismatches 3;
Indels 0; Gaps 0;	Indels 0; Gaps 0;
Qy 1 RPPIKPSWPKKKR 15	Qy 1 RPPIKPSWPKKKR 15
Db 29 RPPIKPSWPKKKR 43	Db 72 RPSPSPSPFAPSRIR 86
RESULT 7	RESULT 9
Q99346 ID Q99326 PRELIMINARY; PRT; 88 AA.	Q7XC25 ID Q7XC25 PRELIMINARY; PRT; 467 AA.
AC Q99326; DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q7XC25; DT 01-OCT-2003 (TREMBLrel. 25, Created)
RT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	RT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative E4 protein.	DE Putative phosphatidylinositol-4-phosphate 5-kinase.
E4.	GN OSJNBA0027P10/22.
OS Human papillomavirus type 82.	OS Oryza sativa (Japonica cultivar-group).
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;	OC Bacteriota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Papillomavirus.	OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OX NCBI_TaxID=129724;	OC Ehrhartoideae; Orzyeae; Oryza.
RN [1] -	NCBI_TaxID=39447;
RP Sequence FROM N.A.	RN
RA Terri M., Burk R.D.;	RP SEQUENCE FROM N.A.
RT "Cervical HPVs in Evolution; Genomic Sequence of IS39/AE2, a subtype of Oncogenic HPV 82 (W13B)." ;	RC STRAIN=cv. Nipponbare;
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.	RA The Rice Chromosome 10 Sequencing Consortium;
RL DR EMBL; AF33961; AAC28453; 1.	RT "In-depth view of structure, activity, and evolution of rice chromosome 10";
DR InterPro; IPR003861; Papilloma_E4.	RT Science 300:1566-1569 (2003).
PFam; PF02711; Pap_E4; 1.	RN [2]
SQ SEQUENCE 88 AA; 10084 MW; 6752D8CF3A9475D7 CRC64;	RP SEQUENCE FROM N.A.
Query Match Score 74.7%; DB 12; Length 88;	RC STRAIN=cv. Nipponbare;
Best Local Similarity 80.0%; Pred. No. 0.012;	RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.,
Matches 12; Conservative 0; Mismatches 3;	RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
Indels 0; Gaps 0;	DR EMBL; AE017120; AAP5505/0.1; -.
Qy 1 RPPIKPSWPKKKR 15	Qy 467 AA; 50906 MW; 8FB85BCC9B980D73 CRC64;
Db 24 RPPIKPSWPKKKR 38	Db 72 RPSPSPSPFAPSRIR 86
RESULT 8	RESULT 10
Q9AY37 ID Q9AY37 PRELIMINARY; PRT; 467 AA.	Q9PZWB ID Q9PZWB PRELIMINARY; PRT; 1242 AA.
AC Q9AY37; DT 01-JUN-2001 (TREMBLrel. 17, Created)	AC Q9PZWB; DT 01-MAY-2000 (TREMBLrel. 13, Created)
RT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)	RT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Putative phosphatidylinositol-4-phosphate 5-kinase.	DE Structural polyprotein.
OS Oryza sativa (Rice).	OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis virus).
OC Bacteriota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	RC
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;	
OC Ehrhartoideae; Orzyeae; Oryza.	
OX NCBI_TaxID=4530;	
RN [1] -	
RP Sequence FROM N.A.	
RC STRAIN=cv. Nipponbare;	

OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;	DR	InterPro; IPR003003; Cys_Ser trypsin.
OC	Alphavirus.	DR	InterPro; IPR003336; Flavi_glycoprote.
RN	[1] - TaxID:11021;	DR	InterPro; IPR007110; Ig-like.
SEQUENCE FROM N.A.		DR	InterPro; IPR006930; Peptidase_S3.
STRAIN-BR56_Bear5122;		DR	PFam; PF00944; Alpha_core_1.
Brault, A.C., Powers A.M., Kang W., Tesh R.B., Shope R.E., Weaver S.C.;		DR	PFam; PF01589; Alpha_E1_glycop_1.
RA		DR	PFam; PF00943; Alpha_E2_glycop_1.
"Genetic and Antigenic Diversity among Eastern and South America";		DR	PRINTS; PF01563; Alpha_E3_glycop_1.
RT		DR	PRINTS; PRO0798; TOGAVIRIN.
viruses from North, Central and South America";		KW	Polyprotein; PRO0798; TOGAVIRIN.
RL	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.	SQ	SEQUENCE 1242 AA; 137434 MW; 2759F20F690BA13 CRC64;
DR	HSSP; P03315; ICP.	Query Match	61.5%; Score 56; DB 12; Length 1242;
DR	GO; GO:0016020; C:membrane; IEA.	Best Local Similarity	66.7%; Pred. No. 9;
DR	GO; GO:0019028; C:viral envelope; IEA.	Matches	10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
DR	GO; GO:0004252; F:serine-type endopeptidase activity; IEA.	Qy	1 RPIPKPSWPWKHR 15
DR	GO; GO:0005198; F:structural molecule activity; IEA.	Db	70 KPAKPKPAAPKKR 84
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR00548; Alpha_E1_glycop.		
DR	InterPro; IPR00533; Alpha_E2_glycop.		
DR	InterPro; IPR00533; Alpha_E3_glycop.		
DR	InterPro; IPR003003; Cys_Ser trypsin.	RESULT 12	
DR	InterPro; IPR00336; Flavi_glycoprote.	ID	Q9PZW6
DR	InterPro; IPR007110; Ig-like.	PRELIMINARY;	PRT; 1242 AA.
DR	InterPro; IPR00930; Peptidase_S3.	AC	Q9PZW6,
DR	PFam; PF00944; Alpha_core_1.	DT	01-MAY-2000 (TRIMBLrel. 13, Created)
DR	PFam; PF01589; Alpha_E1_glycop_1.	DT	01-MAY-2000 (TRIMBLrel. 13, Last sequence update)
DR	PFam; PF00943; Alpha_E2_glycop_1.	DT	01-OCT-2003 (TRIMBLrel. 25, Last annotation update)
DR	PFam; PF01563; Alpha_E3_glycop_1.	DE	Structural polyprotein.
DR	PRINTS; PRO0798; TOGAVIRIN.	OS	Eastern equine encephalitis virus (Eastern equine encephalomyelitis virus).
KW	Polyprotein.	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae; Alphavirus.
SQ	SEQUENCE 1242 AA; 137436 MW; 25A0990E69D16636 CRC64;	OC	NCBI_TAXID-11021;
Query Match	61.5%; Score 56; DB 12; Length 1242;	OX	RN [1]
Best Local Similarity	66.7%; Pred. No. 9;	RP	SEQUENCE FROM N.A.
Matches	10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	RC	STRAIN-BRB5-436087;
Qy	1 RPIPKPSWPWKHR 15	RA	Brault A.C., Powers A.M., Kang W., Tesh R.B., Shope R.E., Weaver S.C.; "Genetic and Antigenic Diversity among Eastern Equine encephalitis viruses from North, Central and South America";
Db	70 KPAKPKPAAPKKR 84	RT	Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
		RL	EMBL; AF15561; AACF04803.1; -.
		DR	HSSP; P03315; IVCP.
		DR	GC; GO:0016020; C:membrane; IEA.
		DR	GC; GO:0019028; C:viral capsid; IEA.
		DR	GC; GO:0019031; C:viral envelope; IEA.
		DR	GC; GO:0004252; F:serine-type endopeptidase activity; IEA.
		DR	GC; GO:0005198; F:structural molecule activity; IEA.
		DR	GC; GO:0006508; P:proteolysis and peptidolysis; IEA.
		DR	InterPro; IPR00548; Alpha_E1_glycop.
		DR	InterPro; IPR00533; Alpha_E2_glycop.
		DR	InterPro; IPR00533; Alpha_E3_glycop.
		DR	InterPro; IPR007110; Ig-like.
		DR	Polyprotein; IPR003006; Ig_MHC.
		DR	InterPro; IPR00930; Peptidase_S3.
		DR	PFam; PF00944; Alpha_core_1.
		DR	PFam; PF01589; Alpha_E1_glycop_1.
		DR	PFam; PF00943; Alpha_E2_glycop_1.
		DR	PFam; PF01563; Alpha_E3_glycop_1.
		DR	PRINTS; PRO0798; TOGAVIRIN.
		DR	PROSITE; PS00290; Ig_MHC.
		KW	Polyprotein.
		SQ	SEQUENCE 1242 AA; 137405 MW; 7254B6D40ED65C3 CRC64;
		Query Match	61.5%; Score 56; DB 12; Length 1242;
		Best Local Similarity	66.7%; Pred. No. 9;
		Matches	10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy	1 RPIPKPSWPWKHR 15	Db	70 KPAKPKPAAPKKR 84

RESULT 13									
Q49870		PRELIMINARY;		PRT;		330 AA.			
ID	049870;								
AC									
DT	01-JUN-1998	(TREMBLrel.	06,	Created)					
DT	01-JUN-1998	(TREMBLrel.	06,	Last sequence update)					
DT	01-OCT-2003	(TREMBLrel.	25,	Last annotation update)					
DE									
GN									
EX1									
OS	Hordium vulgare (Barley).	Streptophytai; Embryophytai; Tracheophytai;							
Eukaryota; Viridiplantae; Spermatophytai; Magnoliophytai; Liliopsida; Poaceae; Pooideae;									
OC									
OC	Triticeae; Hordeum.								
NCBI_TaxID	4513;								
[1]									
RN									
RP									
RA	Doan D.N.P., Sturaro M., Olsen O.A.;	Characterization of a nucleolar cDNA encoding a putative extensin from developing barley grains (Hordeum vulgare L.).	"						
RT									
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.								
DR	Z98204; CAB10894.1; -.								
PIR	T05717; T05717.								
FT	NON_TER 1 1								
SEQUENCE	330 AA;	34218 MW;	E441C3216470A749 CRC64;						
Query Match		59.3%;	Score 54;	DB 10;	Length 330;				
Best Local Similarity	75.0%;	Pred. No. 4.7;							
Matches	9;	Conservative	2;	Mismatches	1;	Indels	0;	Gaps	0;
SQ									
QY		1 RPIPKSPWAK 12							
Db	200 KPVPKSPSPAK 211								
RESULT 14									
Q9LIP1									
ID	Q9LIP1;	PRELIMINARY;							
AC									
DT	01-OCT-2000	(TREMBLrel.	15,	Created)					
DT	01-JUN-2003	(TREMBLrel.	24,	Last sequence update)					
DT	01-JUN-2003	(TREMBLrel.	24,	Last annotation update)					
DE									
OS	Arabidopsis thaliana (Mouse-ear cress).								
Eukaryota; Viridiplantae; Streptophytai; Embryophytai; Tracheophytai;									
Spermatophytai; Magnoliophytai; eudicotyledons; core eudicots; rosids;									
OC									
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.								
NCBI_TaxID	=3702;								
RN									
RP									
RC									
SEQUENCE FROM N.A.									
RC	STRAIN=Columbia;								
RX	MEDLINE=20363039; PubMed=10907853;								
RA	Nakamura Y.;								
RT	"Structural analysis of Arabidopsis thaliana chromosome 3. II.								
RT	Sequence features of the regions of 4,251,695 bp covered by ninety p1,								
RT	TAC and BAC clones";								
PL	DNA Res. 7:217-221 (2000).								
DR	EMBL; AP001298; BAB0195.1; -.								
DR	GO; GO:0005622; C:intracellular; IEA.								
DR	GO; GO:0005840; C:ribosome; IEA.								
DR	GO; GO:0003735; F:structural constituent of ribosome; IEA.								
DR	GO; GO:000412; P:protein biosynthesis; IEA.								
DR	InterPro; IPR001865; Ribosomal S2_1.								
DR	PROSITE; PS00967; Ribosomal S2_1.								
SEQUENCE	356 AA;	39906 MW;	_124A52ADE8C08EF CRC64;						
Query Match		59.3%;	Score 54;	DB 10;	Length 356;				

GenCore version 5.1.6	May 27, 2004, 16:03:03 ;	US-10-008-5244A-168	44
Copyright (c) 1993 - 2004 Compugen Ltd.	search time 33:2308 seconds (without alignments)	75	44
protein - protein search, using sw model	102.031 Million cell updates/sec	1 PRESFWAKKHR 12	44
on:	US-10-008-5244A-168	BLOSUM62	45
title:	75	Gapext 0.5	45
rfc score:	1	1586107 seqs, 282547505 residues searched.	45
quence:			
oring table:			

minimum DB seq length: 0
maximum DB seq length: 2000000000

st-processing: Minimum Match 0%
Listning First 45 summaries
Maximum Match 100%
RESULT 1
AAW62279
TT AAW62279 standard: dendrite: 12 AA

```

database : A_Geneseq_29Jan04:*
 1: geneseqP1980s:*
 2: geneseqP1990s:*
 3: geneseqP2000s:*
 4: GeneseqP2001s:*
 5: GeneseqP2002s:*
 6: geneseqP2003as:*
 7: geneseqP2003bs:*
 8: GeneseqP2004s:*
AAMW62279;
AC XX
DT 24-SEP-1998 (first entry)
XX
DB HPV16 hydrophilic region peptide #3.
XX
KW HPV16; human papilloma virus; epithelial tumour; cervical cancer;
KW precancerous cervical lesion; screening; detection; infection; cervix;
KW HPV E4.

```

No. is the number of results predicted by chance to have a greater score than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
AAW62279
ID AAW62279 standard; peptide; 12 AA.

AC	AAW2279;
XX	
DT	24-SEP-1998 (first entry)
XX	
DE	HPV16 hydrophilic region peptide #3.
XX	
KW	HPV16; human papilloma virus; epithelial
KW	precancerous cervical lesion; screen
KW	HPV E4

OS Human papillomavirus.
XX WO9825145-A1.
PN

PD 11-JUN-1998 .
 XX
 PF 03-DEC-1997; 97WO-GB003321.
 XX
 PR 03-DEC-1996; 96GB-00025142.
 PR 05-SEP-1997; 97GB-00018745.
 XX
 PA (MEDI-) MEDICAL RES COUNCIL.
 XX
 PI Doorbar J;
 XX
 DR WPI: 1998-333497/29 .

PT Detecting papilloma virus infection using molecule binding to E4 protein
PT - useful, e.g. in screening for pre-cancerous cervical lesions and to
PT determine type(s) of human papilloma virus infecting human patients.

XX

PS Claim 11; Page 37; 52pp; English.

XX A new method has been developed for detecting a papilloma virus infection
CC in an organism. The method comprises: (i) obtaining a sample of cells
CC from the potential infection site; (ii) contacting the sample with a
CC molecule binding specifically to papilloma virus E4 protein, and (iii)
CC monitoring the binding. The method is useful to detect papilloma virus
CC infections in organisms (especially mammals) and especially HPV
CC infections (e.g. with HPV16, 18, 33, 35, 45, 51, 56, 58 or 61) in humans.
CC Papilloma viruses cause epithelial tumours in humans varying in severity
CC depending on the infection site and HPV type involved. The method is
CC particularly useful to determine papilloma infection in the mammalian
CC

humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and hence clinical prognosis) is dependent on HPV type. Molecules capable of binding E4 are also useful to target anticancer/antiviral agents capable of destroying papilloma viruses and/or papilloma virus-infected cells.

The present sequence represents a specifically claimed HPV16 peptide found in the hydrophilic region

XX

Sequence 12 AA;

Query Match 100.0%; Score 75; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKKHR 12
Dy 1 SKPSPWPKKHR 12

RESULT 2

AAW62278

ID AAW62278 standard; peptide; 15 AA.

XX

AC AAW62278;

XX

DT 24-SEP-1998 (first entry)

XX

DE HPV16 hydrophilic region Peptide #2

XX

HPV16; human papilloma virus; epithelial tumour; cervical cancer; precancerous cervical lesion; screening; detection; infection; cervix; HPV E4.

XX

OS Human papillomavirus.

XX

EN WO9825145-A1 .

XX

DD 11-JUN-1998.

XX

FP 03-DEC-1997; 97WO-GB003321.

XX

PR 03-DEC-1996; 96GB-00025142.

XX

PR 05-SEP-1997; 97GB-00018745.

XX

PA (MEDI-) MEDICAL RES COUNCIL.

XX

FI Doorbar J;

XX

DR WPI; 1998-333497/29.

XX

Detecting papilloma virus infection using molecule binding to E4 protein - useful, e.g. in screening for pre-cancerous cervical lesions and to determine type(s) of human papilloma virus infecting human patients.

XX

FS Claim 10; Page 37; 52pp; English.

XX

A new method has been developed for detecting a papilloma virus infection in an organism. The method comprises: (i) obtaining a sample of cells from the potential infection site; (ii) contacting the cells with a molecule binding specifically to papilloma virus E4 protein, and (iii) monitoring the binding. The method is useful to detect papilloma virus infections in organisms (especially mammals) and especially HPV

Papilloma viruses cause epithelial tumours in humans varying in severity depending on the infection site and HPV type involved. The method is particularly useful to determine papilloma infection in the mammalian cervix and especially to screen for pre-cancerous cervical lesions in humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and

hence clinical prognosis) is dependent on HPV type. Molecules capable of binding E4 are also useful to target anticancer/antiviral agents capable of destroying papilloma virus and/or papilloma virus-infected cells. The present sequence represents a specifically claimed HPV16 peptide found in the hydrophilic region

XX Sequence 15 AA;

Query Match 100.0%; Score 75; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00039;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

SQ

RESULT 3
AAR15562
ID AAR15562 standard; protein; 20 AA.
XX
AC AAR15562;
XX DT 02-MAR-1992 (first entry)
XX DE Immunopeptide #2 derived from HPV16 E4 peptide.
XX cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
XX squamous cell carcinoma; ELISA; HPV 16.
OS Synthetic.
PN WO9118294-A.
XX PR 11-MAY-1990; 90SE-00001705.
XX PR 11-MAY-1990; 90SE-00001705.
XX PR 28-NOV-1991.
XX PR 11-MAY-1990; 90SE-00001705.
XX PR 11-MAY-1990; 90SE-00001705.
XX PA (MEDS-) MEDSCAND AB.
XX PI Dillner J, Dillner L, Cheng HM;
XX DR WPI; 1991-369390/50.
XX PT Diagnosis of human papilloma virus infection and PV-carrying tumours - using synthetic peptide(s) to detect virus specific antigen-antibody complexes by immunoassay.
XX Disclosure: Page 38; 72pp; English.

This is one of a large number of peptides which have been synthesised on the basis of the amino acid sequences for the E2, E4, E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The selection of peptide sequences was based on the assumption that immunoreactive region might be situated in the same relative region of a protein from different HPV types. The peptides were used in diagnostic immunoassays to detect HPV-infection. See AAR15523-R15601

XX Sequence 20 AA;

Query Match 100.0%; Score 75; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00052;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKKHR 12
Db 2 PKPSPWPKKHR 13

RESULT 4
AAW62277

ID	AAW62277 standard; peptide; 26 AA.	DT	21-MAY-2002 (first entry)
AC	AAW62277;	XX	Hydrophilic region of HPV16 E4 protein.
XX	24-SEP-1998 (first entry)	DT	Papilloma virus associated antigen; cell proliferation marker; cervical malignancy; human papillomavirus infection; HPV; lesion; cellular abnormality; cellular proliferation; cellular growth; dysplasia; neoplasia; cancer; Papilloma smear test; E4 protein.
DE	HPV16 hydrophilic region peptide #1.	XX	Human papillomavirus type 16.
XX	HPV16; human papilloma virus; epithelial tumour; cervical cancer; preancerous cervical lesion; screening; detection; infection; cervix; HPV E4.	XX	OS
XX	Human papillomavirus.	XX	XX
OS	W09825145-A1.	XX	PN WO200208764-A1.
PN	W09825145-A1.	PD	PD 31-JAN-2002.
XX	11-JUN-1998.	PD	16-MAR-2001; 2001WO-GB001176.
XX	03-DEC-1997; 97W0-GB003321.	PR	24-JUL-2000; 2000GB-00018140.
PF	03-DEC-1996; 96GB-00025142.	XX	XX
PR	05-SEP-1997; 97GB-00018745.	PA	(MEDI-) MEDICAL RES COUNCIL.
XX	(MEDI-) MEDICAL RES COUNCIL.	XX	Doorbar J;
PA	Doorbar J;	PI	Doorbar J;
XX	DR WPI; 2002-188648/24.	XX	DR WPI; 2002-188648/24.
XX	Detecting abnormalities e.g. abnormal cellular proliferation, in a sample from a patient comprises contacting cells with a molecule which can bind a papilloma virus associated antigen, or a cell proliferation or viral activity marker.	XX	Detecting abnormalities e.g. abnormal cellular proliferation, in a sample from a patient relates to a method for detecting abnormalities in a sample from a patient. The method comprises contacting a sample of the patient's cells with two or more molecules, where at least one molecule is capable of binding a papilloma virus associated antigen, and at least one molecule is capable of binding a cell proliferation marker. The method is useful for simultaneously screening for abnormalities which may or can lead to cervical malignancy, for human papillomavirus (HPV) infections, and precursor lesions or other conditions which occur with cervical malignancy. The method is also useful for assessing the risk associated with cellular abnormality in a patient sample, and for determining, assessing or diagnosing the presence or absence of abnormal cellular proliferation, cellular growth abnormality, dysplasia, neoplasia, or a pre-cancerous or cancerous state in a tissue. The new method is much simpler, and yields more information more quickly than conventional papilloma smear testing programmes. Compared with previous methods of screening, the new method has reduced chances of false negatives occurring, requires fewer samples to gain the same amount of information, and alleviates the need for repeated or further testing. The present sequence represents the hydrophilic region of the HPV16 E4 protein and alleviates the need for repeated or further testing. The present sequence represents the hydrophilic region of the HPV16 E4 peptide.
PS	PS Disclosure; Page 23; 90pp; English.	XX	Sequence 26 AA;
PT	PT	CC	Query Match 100.0%; Score 75; DB 5; Length 26;
PT	PT	CC	Best Local Similarity 100.0%; Pred. No. 0.00068; Mismatches 0; Indels 0; Gaps 0;
PT	PT	CC	Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
PT	PT	CC	XX
PS	PS Disclosure; Page 23; 90pp; English.	XX	XX
XX	Detecting papilloma virus infection using molecule binding to E4 protein - useful, e.g. in screening for pre-cancerous cervical lesions and to determine type(s) of human papilloma virus infecting human patients.	XX	XX
XX	Claim 9; Page 37; 52pp; English.	XX	XX
XX	A new method has been developed for detecting a papilloma virus infection in an organism. The method comprises: (i) obtaining a sample of cells from the potential infection site; (ii) contacting the cells with a molecule binding specifically to papilloma virus E4 protein, and (iii) monitoring the binding. The method is useful to detect papilloma virus infections in organisms (especially mammals) and especially HPV infections (e.g. with HPV16, 18, 33, 35, 45, 51, 56, 58 or 61) in humans. Papilloma viruses cause epithelial tumours in humans varying in severity depending on the infection site and HPV type involved. The method is particularly useful to determine papilloma infection in the mammalian cervix and especially to screen for pre-cancerous cervical lesions in humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and hence clinical prognosis) is dependent on HPV type. Molecules capable of binding E4 are also useful to target anti-cancer/antiviral agents capable of destroying papilloma viruses and/or papilloma virus-infected cells. The present sequence represents a specifically claimed HPV16 peptide found in the hydrophilic region	XX	XX
XX	Sequence 26 AA;	XX	XX
Qy	Query Match, Best Local Similarity, Matches	XX	XX
Db	DB 5; Length 26; Score 75; Pred. No. 0.00068; Mismatches 0; Indels 0; Gaps 0;	XX	XX
Qy	1 PKPSPWAPKKHR 12	1 PKPSPWAPKKHR 12	1 PKPSPWAPKKHR 12
Db	4 PKFSPPWAPKKHR 15	4 PKFSPPWAPKKHR 15	4 PKFSPPWAPKKHR 15
RESULT 6	AAR15564	RESULT 6	AAR15564
AAU75260	ID AAR15564 standard; peptide; 26 AA.	AAU75260	ID AAR15564 standard; peptide; 30 AA.
XX	AC AAR15564;	XX	AC AAR15564;
XX	DT 02-MAR-1992 (first entry)	XX	DT 02-MAR-1992 (first entry)
AC	DB Immunopeptide #4 derived from HPV16 E4 peptide.	AC	DB Immunopeptide #4 derived from HPV16 E4 peptide.

XX cervical cancer; cervical intraepithelial neoplasia; CIN; wart;
 KW squamous cell carcinoma; ELISA; HPV 16.
 XX Synthetic.
 OS WO9118294 A.
 PN XX
 PD 28-NOV-1991.
 XX PFF 11-MAY-1990; 90SE-00001705.
 PR 11-MAY-1990; 90SE-00001705.
 PA (MEDS-) MEDSCAND AB.
 XX Dillner J, Dillner L, Cheng HM;
 XX WPI; 1991-369390/50.
 DR XX

PT Diagnosis of human papilloma virus infection and PV-carrying tumours -
 PT using synthetic peptide(s) to detect virus specific antigen antibody
 PT complexes by immunassay.
 XX Disclosure; Page 38; 72PP; English.

CC This is one of a large number of peptides which have been synthesised on
 CC the basis of the amino acid sequences for the E2, E4, E7, L1 or L2
 CC proteins of HPV 1, 5, 6, 8, 11, 16, 31 and 33. The selection of
 CC peptide sequences was based on the assumption that an immunoreactive
 CC region might be situated in the same relative region of a protein from
 CC different HPV types. The peptides were used in diagnostic immunoassays to
 CC detect HPV-infection. See AARI5523-R15601.
 XX SQ Sequence 30 AA;

Query Match Score 75; DB 2; Length 30;
 Best Local Similarity 100.0%; Pred. No. 0.00078;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPKKER 12
 Db 10 PKPSPWAPKKER 21

RESULT 7
 AAR-4288
 ID AAR14288 standard; protein; 14 AA.
 AC AARI14288;
 XX DT 02-JAN-1992 (first entry)
 DE Seroreactive epitope #2 of HPV16 protein E4.
 KW HPV16-dependent human disease; E6; E7; L1.
 XX OS Synthetic.

XX PFF 02-JAN-1992 (first entry)
 DE Seroreactive epitope #2 of HPV16 protein E4.
 KW HPV16-dependent human disease; E6; E7; L1.
 XX OS Synthetic.

XX PT Seroreactive epitope(s) of human papilloma-virus 16 proteins - for use
 vaccines and diagnosis.
 XX Claim 1; Page 11; 15pp; English.
 OS This is one of two seroreactive epitopes identified from HPV16 protein
 CC E4. Peptides containing this epitope are also claimed and can be used to
 CC generate antibodies to HPV. See AARI14287-R14302 and AAQ14168-Q14171.
 XX SQ Sequence 14 AA;
 Query Match Score 68; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 0.0036;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KPSPWAPKKER 12
 Db 1 KPSPWAPKKER 11
 RESULT 8
 AARI14287
 ID AAR14287 standard; peptide; 11 AA.
 XX AC AARI14287;
 XX DT 02-JAN-1992 (first entry)
 DE Seroreactive epitope #1 of HPV16 protein E4.
 KW HPV16-dependent human disease; E6; E7; L1.
 XX OS Synthetic.
 XX PN EP451550-A.
 XX PD 16-OCT-1991.
 XX PF 19-MAR-1991; 91EP-00101197.
 XX PR 20-MAR-1990; 90EP-00105222.
 XX PA (BEHRINGWERKE AG.
 XX PI Muller M, Giessmann L;
 XX DR WPI; 1991-304643/42.
 XX PT Sero-active epitope(s) of human papilloma-virus 16 proteins - for use
 vaccines and diagnosis.
 XX SQ Sequence 11 AA;

Query Match Score 82.7%; DB 2; Length 11;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPKK 10
 Db 2 PKPSPWAPKK 11
 RESULT 9
 AA003023
 ID AA003023 standard; protein; 105 AA.
 XX AC AA003023;

PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US004927.
 XX
 PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 WPI; 2001-514838/56.
 DR N-PSDB; AAI84881.
 XX
 Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

Claim 20; SEQ ID NO 18042; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79841-AAI93841) and the encoded proteins (AA00010-AA013910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

Sequence 115 AA;
 SQ Query Match Score 50; DB 4; Length 115;
 Best Local Similarity 72.7%; Pred. No. 11;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PKPSPWPKKH 11
 | | | | | | :
 91 PPPPPWPKKN 101

Result 12
 ABG08384
 ID ABG08384 standard; protein; 72 AA.
 XX
 AC ABG08384;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #8575.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649367.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362-A2.
 DR N-PSDB; AAS7771.
 XX New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics. Gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.
 XX
 PS Claim 20; SEQ ID NO 38943; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, Gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AIG00010-AIG30377 represent novel human diagnostic amino acid sequences. The sequence data for this patient did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp://wipo.int/pub/published_pct_sequences

Sequence 72 AA;
 SQ Query Match Score 46; DB 4; Length 72;
 Best Local Similarity 60.0%; Pred. No. 25;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 PSPWPKKH 12
 :| | | | :
 Db 20 PTPWAPVQHK 29

RESULT 13
 ABR41361
 ID ABR41361 standard; protein; 91 AA.
 XX
 AC ABR41361;
 XX
 DT 02-JUN-2003 (first entry)
 XX Human DITHP transcription factor.
 DB
 XX Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;
 KW cancer; cell proliferative disorder; autoimmune disorder;
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;
 KW neurological disorder; gastrointestinal disorder; transport disorder;
 KW connective tissue disorder; drug screening; proteome analysis;
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;
 KW disease model; toxicological testing; transcript imaging;
 KW
 XX Homo sapiens.
 XX
 PN WO200297031-A2.
 XX
 PR 05-DEC-2002.
 XX
 PR 27-MAR-2003; 2002WO-US010056.
 XX
 PR 28-MAR-2001; 2001US-0279619P.
 PR 29-MAR-2001; 2001US-0280067P.

PR	29-MAR-2001; 2001US-0280068P.	XX	Human secreted protein, SEQ ID NO: 6028.	
PR	16-MAY-2001; 2001US-0291180P.	DE		
PR	17-MAY-2001; 2001US-0291180P.	XX		
PR	17-MAY-2001; 2001US-0291184P.	KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation; gene therapy; chromosome mapping.	
PR	19-JUN-2001; 2001US-02994128P.	KW		
PR	20-JUN-2001; 2001US-02994128P.	KW		
PR	20-JUN-2001; 2001US-02994128P.	XX	Homo sapiens.	
PA	(INCY-) INCYTE GENOMICS INC.	OS		
PA	PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;	XX		
PA	Dufour GE, Hillman JL, Yu JY, Nason O, Yap PB, Amsthey SR;	XX		
PA	Daugerty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Garstin EH;	XX		
PA	Peralta CH, Davida MH, Lewis SA, Chen AJ, Panzer SR, Harris B;	XX		
PA	Flores V, Marwaha R, Lo A, Ian YY, Urashka ME;	XX		
WPI	DR 2003-129518/12.	XX		
N-PSDB; ACC46301.	DR 2003-129518/12.	PA	(GEST) GENSET.	
XX	PT Novel human diagnostic and therapeutic polypeptide useful for identifying test compound which specifically binds to a polypeptide encoded by human diagnostic and therapeutic polynucleotide, and to induce antibodies.	XX		
PS	Claim 27; SEQ ID NO 896; 591pp; English.	PT Novel human diagnostic and therapeutic polynucleotides designated dithp (ACCG080-ACC46749) and to their encoded proteins (DITHP; APR1136-APR1812). The invention also relates to polynucleotide sequences at least 90% identical to the dithp cDNA sequences of the invention; recombinant vectors, host cells and transgenic organisms comprising a dithp nucleic acid sequence; the recombinant proteins specific for DITHP proteins; antibodies specific for DITHP proteins; microarrays comprising dithp nucleic acid sequences; methods of detecting dithp nucleotide and protein sequences; methods of screening for compounds which specifically bind a DITHP protein; and methods of assessing the toxicity of test compounds using a dithp hybridisation probe. Dithp nucleic acid sequences and DITHP proteins may be used in the diagnosis of a wide variety of conditions including cancer and other cell proliferative disorders; autoimmune or inflammatory disorders; bacterial, viral, fungal or parasitic infections; hormonal disorders; metabolic disorders; neurological disorders; gastrointestinal disorders; transport disorders; and connective tissue disorders. They may also be used to screen for modulators of protein activity or gene expression. DITHP proteins can additionally be used in analysis of the proteome of a tissue or cell type and to induce antibodies. The dithp nucleic acids are additionally useful in somatic or germline gene therapy of the disorders mentioned above, as a source of antisense sequences, as a source of probes and primers, in genotyping and identification of individuals, in the generation of transgenic animal models of human disease or knock in humanised animals, in toxicological testing, and in transcript imaging. The present sequence represents a DITHP protein which has transcription factor activity. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	XX	Sequence 91 AA;
PS	XX	Query Match Score 46; DB 6; Length 91; Best Local Similarity 61.3%; Pred. No. 32; Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	XX	RESULT 15
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 61.3%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	ABU05501
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	ABU05501;
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	DT 08-APR-2003 (first entry)
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	M. tuberculosis and M. leprae marker protein #152.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	KW Mycobacterioses; survival; virulence; protective antigen; vaccine; mycobacterial disease; tuberculosis; leprosy.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	OS Mycobacterium tuberculosis.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	OS Mycobacterium leprae.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	PN WO200274903-A2.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	PD 26-SEP-2002.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	PRERESULT 14
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	AAG01947
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	AAG01947 standard; protein; 121 AA.
PS	XX	Query Match Score 46; DB 3; Length 121; Best Local Similarity 54.5%; Pred. No. 43; Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	XX	DT 06-OCT-2000 (first entry)

XX 22-FEB-2002; 2002WO-IB001973.
 XX
 PR 22-FEB-2001; 2001US-0270123P.
 XX (INSP) INST PASTEUR.
 XX
 PI Cole S;
 XX
 DR WPI; 2002-759885/82.
 XX
 PT Identifying and selecting genes for survival or virulence of mycobacteria
 PT by a comparative genomic analysis of the sequences of Mycobacterium
 PT tuberculosis and M. leprae.

XX
 PS Claim 17; Page 322-323; 874pp; English.

XX This invention relates to a novel method for identifying essential genes
 CC for survival or virulence of mycobacteria species. The method comprises
 CC aligning the genomic sequence of a first mycobacterium species on a
 CC genomic sequence of a second mycobacterium species and selecting a
 CC polynucleotide sequence that is highly conserved in both genomes with no
 CC counterparts in other bacterial genomes and that corresponds to
 CC an essential gene for the survival or virulence of mycobacterium species.
 CC The method of the invention is useful for detecting M. tuberculosis or M.
 CC leprae infection. The method reduces the number of potential new targets
 CC and protective antigens for new drugs and vaccine compositions to treat
 CC and prevent mycobacterial diseases, particularly tuberculosis and
 CC leprosy. The present sequence represents a marker protein from
 CC Mycobacterium tuberculosis and Mycobacterium leprae identified using the
 CC method of the invention

XX Sequence 230 AA;

Query	Match	Score	DB	Length
Best	Local Similarity	46;	5;	230;
Matches	Pred. No.	77.8%;	81;	
	Mismatches	1;	Indels	0;
			Gaps	0;

QY 1 PKPSPWAPK 9
 Db ||||| |||:
 110 PKPSKWPAPR 118

Search completed: May 27, 2004, 16:18:30
 Job time : 34.2308 secs

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:15:59 ; Search time 9.69231 Seconds
 (Without alignments)
 63.918 Million cell updates/sec

Title: US-10-008-524A-168

Perfect score: 75 Sequence: 1 PKPSPWAPKHR 12

Scoring table: BLOSUM62 Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database :	Issued Patents AA:*
	1: /cgn2_6/potodata/2/iaa/5A_COMB.pep:*
	2: /cgn2_6/potodata/2/iaa/5B_COMB.pep:*
	3: /cgn2_6/potodata/2/iaa/6A_COMB.pep:*
	4: /cgn2_6/potodata/2/iaa/6B_COMB.pep:*
	5: /cgn2_6/potodata/2/iaa/PCTUSCOMB.pep:*
	6: /cgn2_6/potodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

%
SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	75	100.0	12	4	US-09-314-268-168 Sequence 168, APP
2	75	100.0	15	4	US-09-314-268-167 Sequence 167, APP
3	75	100.0	20	2	US-08-934-915-40 Sequence 40, APP
4	75	100.0	20	2	US-08-934-915-170 Sequence 4, APP
5	75	100.0	25	4	US-09-314-268-4 Sequence 42, APP
6	75	100.0	30	2	US-08-934-915-42 Sequence 33, APP
7	75	100.0	92	4	US-09-314-268-37 Sequence 117, APP
8	68	90.7	35	4	US-09-314-268-33 Sequence 36, APP
9	52	69.3	8	4	US-09-314-268-33 Sequence 32, APP
10	52	69.3	8	4	US-09-314-268-37 Sequence 118, APP
11	51	68.0	8	4	US-09-314-268-36 Sequence 38, APP
12	51	68.0	34	4	US-09-314-268-118 Sequence 119, APP
13	50	66.7	8	4	US-09-314-268-34 Sequence 35, APP
14	50	66.7	8	4	US-09-314-268-35 Sequence 21, APP
15	49	65.3	325	4	US-09-921-099A-21 Sequence 213, APP
16	48	64.0	143	4	US-09-254-991A-21367 Sequence 910, APP
17	45	60.0	8	4	US-09-314-268-32 Sequence 12, APP
18	45	60.0	8	4	US-09-314-268-38 Sequence 12, APP
19	45	60.0	39	4	US-09-314-268-119 Sequence 20580, APP
20	44	58.7	97	4	US-09-282-991A-20580 Sequence 17588, APP
21	44	58.7	299	4	US-09-252-991A-17588 Sequence 910, APP
22	43	57.3	32	4	US-09-205-928-910 Sequence 12, APP
23	43	57.3	524	2	US-08-928-692-32 Sequence 12, APP
24	43	57.3	524	4	US-09-339-972-12 Sequence 12, APP
25	43	57.3	589	4	US-09-654-979A-4 Sequence 4, APP
26	43	57.3	959	4	US-09-232-991A-23758 Sequence 23758, APP
27	42	56.0	40	4	US-09-314-268-106 Sequence 106, APP

ALIGNMENTS

RESULT 1					
US-09-314-268-168					
; Sequence 168, Application US/09314268					
; GENERAL INFORMATION:					
; APPLICANT: Doorbar, John					
; TITLE OF INVENTION: VIRUSES					
; FILE REFERENCE: 3789/80902					
; CURRENT APPLICATION NUMBER: US/09/314,268					
; CURRENT FILING DATE: 1999-03-19					
; EARLIER APPLICATION NUMBER: 09/314,268					
; EARLIER FILING DATE: 1999-05-18					
; NUMBER OF SEQ ID NOS: 179					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO: 168					
; LENGTH: 12					
; TYPE: PRT					
; ORGANISM: Human papillomavirus type 16					
US-09-314-268-168					
Query Match					
Best Local Similarity 100.0%; Score 75; DB 4; Length 12;					
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	PRPSPWAPKHR 12			
Db	1	PRPSPWAPKHR 12			

RESULT 2

US-09-314-268-167					
; Sequence 167, Application US/09314268					
; GENERAL INFORMATION:					
; APPLICANT: Doorbar, John					
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA					
; FILE REFERENCE: 3789/80902					
; CURRENT APPLICATION NUMBER: US/09/314,268					
; CURRENT FILING DATE: 1999-03-19					
; EARLIER APPLICATION NUMBER: 09/314,268					
; EARLIER FILING DATE: 1999-05-18					
; NUMBER OF SEQ ID NOS: 179					
; SOFTWARE: PatentIn Ver. 2.1					
; SEQ ID NO: 167					
; LENGTH: 15					
; TYPE: PRT					
; ORGANISM: Human papillomavirus type 16					
US-09-314-268-167					
Query Match					
Best Local Similarity 100.0%; Score 75; DB 4; Length 12;					
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	PRPSPWAPKHR 12			
Db	1	PRPSPWAPKHR 12			

Query Match 100.0%; Score 75; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.4e-05; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0; Indels 0;

Qy 1 PKPSPWPKHR 12
 Db 4 PKPSPWPKHR 15

RESULT 3
 US-08-934-915-40
 ; Sequence 40, Application US/08934915
 ; GENERAL INFORMATION:
 ; APPLICANT: DILLNER, JOAKIM
 ; APPLICANT: CHENG, HWEE-MING
 ; TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN
 ; PAPILLOMAVIRUS 1, 5, 6, 8,
 ; TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6,
 ; 11, 16, 31, 33 AND 56,
 ; TITLE OF INVENTION: USEFUL IN IMMUNOASSAY FOR
 ; IMMUNOASSAY FOR
 ; NUMBER OF SEQUENCES: 1.93
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MASON & ASSOCIATES, P.A.
 ; STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
 ; CITY: CLEARWATER
 ; STATE: FLORIDA
 ; COUNTRY: U.S.A.
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 3.0
 ; SOFTWARE: Microsoft Word 6.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/934 915
 ; FILING DATE: 22-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/949, 836
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: LOUISE A. FOUTCH
 ; CITY: CLEARWATER
 ; STATE: FLORIDA
 ; COUNTRY: U.S.A.
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: FLOPPY disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 3.0
 ; SOFTWARE: Microsoft Word 6.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/934, 915
 ; FILING DATE: 22-SEP-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; FILING DATE: 07/949, 836
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: LOUISE A. FOUTCH
 ; REGISTRATION NUMBER: 37,133
 ; REFERENCE/DOCKET NUMBER: 1946.6
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 813-538-3800
 ; TELEX: 813-538-3820
 ; INFORMATION FOR SEQ ID NO: 170:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 20 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: Linear
 ; MOLECULE TYPE: peptide
 ; US-08-934 915-170

Query Match 100.0%; Score 75; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 0.00012; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0; Indels 0;

Qy 1 PKPSPWPKHR 12
 Db 2 PKPSPWPKHR 13

RESULT 5
 US-09-314-268-4
 ; Sequence 4, Application US/09314268
 ; Patent No. 6346377
 ; GENERAL INFORMATION:
 ; APPLICANT: DOOBAR, JOHN
 ; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
 ; FILE REFERENCE: 3789/80902
 ; CURRENT APPLICATION NUMBER: US/09/314, 268
 ; CURRENT FILING DATE: 1999-03-19
 ; EARLIER APPLICATION NUMBER: 1999-05-18
 ; NUMBER OF SEQ ID NOS: 179
 ; SEQ ID NO: 4
 ; LENGTH: 25
 ; TYPE: PRT
 ; ORGANISM: Human papillomavirus type 16
 ; US-09-314-268-4

Query Match 100.0%; Score 75; DB 4; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.00012; Gaps 0;
 Matches 12; Conservative 0; Mismatches 0; Indels 0;

Qy 1 PKPSPWPKHR 12
 Db 2 PKPSPWPKHR 13

RESULT 4
 US-08-934-915-170
 ; Sequence 170, Application US/08934915
 ; Patent No. 5931412
 ; GENERAL INFORMATION:
 ; APPLICANT: DILLNER, JOAKIM

Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FILE REFERENCE: 3789/80902
CURRENT APPLICATION NUMBER: US/09/314,268
CURRENT FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: 09/314,268
EARLIER FILING DATE: 1999-05-18
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 90
LENGTH: 92
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-09-314-268-90

RESULT 6

Query Match 100.0%; Score 75; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKHR 12
Db 29 PKPSPWPKHR 40

RESULT 8

Query Match 100.0%; Score 75; DB 4; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKHR 12
Db 29 PKPSPWPKHR 40

RESULT 9

Query Match 90.7%; Score 68; DB 4; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 KSPWPWPKHR 12
Db 1 KSPWPWPKHR 11

RESULT 9

Query Match 100.0%; Score 75; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKHR 12
Db 10 PKPSPWPKHR 21

RESULT 7

Query Match 100.0%; Score 75; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PKPSPWPKHR 12
Db 10 PKPSPWPKHR 21

GENERAL INFORMATION:
Patent No. 6346377
SEQUENCE 42, Application US/08934915
GENERAL INFORMATION:
PATENT NO. 5932412

APPLICANT: DILLNER, JOAKIM
APPLICANT: CHENG, HWEE MING
TITLE OF INVENTION: SYNTHETIC PEPTIDES OF HUMAN PAPILLOMAVIRUS 1, 5, 6, 8, 11, 16, 18, 31, 33 AND 56, USEFUL IN IMMUNOASSAY FOR DIAGNOSTIC PURPOSES
TITLE OF INVENTION: PAPILLOMAVIRUS 1, 5, 6, 8, 11, 16, 18, 31, 33 AND 56, USEFUL IN IMMUNOASSAY FOR DIAGNOSTIC PURPOSES
NUMBER OF SEQUENCES: 193

COUNTRY: U.S.A.
ADDRESSEE: MASON & ASSOCIATES, P.A.
STREET: 17757 U.S. HWY. 19 NORTH, SUITE 500
CITY: CLEARWATER
STATE: FLORIDA

COMPUTER READABLE FORM:
MEDIUM: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: WINDOWS 3.0
SOFTWARE: MICROSOFT WORD 6.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/934,915
FILING DATE: 22-SEP-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/949,836
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: LOUISE A. FOUTCH
REGISTRATION NUMBER: 37,133
TELECOMMUNICATION INFORMATION:
TELEPHONE: 813-538-3800
TELEFAX: 813-538-3820

INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide

US-08-934-915-42

GENERAL INFORMATION:
Patent No. 6346377
SEQUENCE 90, Application US/09314268
GENERAL INFORMATION:
APPLICANT: Doorbar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
TITLE OF INVENTION: VIRUSES
OTHER INFORMATION: Description of Artificial Sequence: synthetic

GENERAL INFORMATION:
Patent No. 6346377
SEQUENCE 90, Application US/09314268
GENERAL INFORMATION:
APPLICANT: Doorbar, John
TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
TITLE OF INVENTION: VIRUSES
OTHER INFORMATION: Description of Artificial Sequence: synthetic

i OTHER INFORMATION: octapeptide antigen
 us-09-314-268-33

Query Match	69.3%;	Score 52;	DB 4;	Length 8;	
Best Local Similarity	100.0%;	Pred. No.	3e+05;		
Matches	8;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		

QY 1 PKPSPWAP 8
Db 1 PKPSPWAP 8

RESULT 10

US-09-314-268-37

i Sequence 37, Application US/09314268

i Patent No. 6346377

i GENERAL INFORMATION:

i APPLICANT: Doorbar, John

i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA

i TITLE OF INVENTION: VIRUSES

i FILE REFERENCE: 3789/80902

i CURRENT APPLICATION NUMBER: US/09/314,268

i CURRENT FILING DATE: 1999-03-19

i EARLIER APPLICATION NUMBER: 09/314,268

i EARLIER FILING DATE: 1999-05-18

i NUMBER OF SEQ ID NOS: 179

i SOFTWARE: PatentIn Ver. 2.1

i SEQ ID NO: 118

i LENGTH: 34

i TYPE: PRT

i ORGANISM: Human papillomavirus type 35

i US-09-314-268-118

Query Match 68.0%; Score 51; DB 4; Length 34;
Best Local Similarity 72.7%; Pred. No. 0.5%;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 KPSWPKHR 12
Db 1 KPSPWPKHR 11

RESULT 13

US-09-314-268-34

i Sequence 34, Application US/09314268

i Patent No. 6346377

i GENERAL INFORMATION:

i APPLICANT: Doorbar, John

i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA

i TITLE OF INVENTION: VIRUSES

i FILE REFERENCE: 3789/80902

i CURRENT APPLICATION NUMBER: US/09/314,268

i CURRENT FILING DATE: 1999-03-19

i EARLIER APPLICATION NUMBER: 09/314,268

i EARLIER FILING DATE: 1999-05-18

i NUMBER OF SEQ ID NOS: 179

i SOFTWARE: PatentIn Ver. 2.1

i SEQ ID NO: 34

i LENGTH: 8

i TYPE: PRT

i ORGANISM: Artificial Sequence

i FEATURE:

i OTHER INFORMATION: Description of Artificial Sequence: synthetic

i OTHER INFORMATION: Octapeptide antigen

i US-09-314-268-37

Query Match 69.3%; Score 52; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 PWAPKHR 12
Db 1 PWAPKHR 8

RESULT 11

US-09-314-268-36

i Sequence 36, Application US/09314268

i Patent No. 6346377

i GENERAL INFORMATION:

i APPLICANT: Doorbar, John

i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA

i TITLE OF INVENTION: VIRUSES

i FILE REFERENCE: 3789/80902

i CURRENT APPLICATION NUMBER: US/09/314,268

i CURRENT FILING DATE: 1999-03-19

i EARLIER APPLICATION NUMBER: 09/314,268

i EARLIER FILING DATE: 1999-05-18

i NUMBER OF SEQ ID NOS: 179

i SOFTWARE: PatentIn Ver. 2.1

i SEQ ID NO: 36

i LENGTH: 8

i TYPE: PRT

i ORGANISM: Artificial Sequence

i FEATURE:

i OTHER INFORMATION: Description of Artificial Sequence: synthetic

i OTHER INFORMATION: Octapeptide antigen

i US-09-314-268-36

Query Match 68.0%; Score 51; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KPSWPK 9
Db 1 KPSPWPK 8

RESULT 14

US-09-314-268-35

i Sequence 35, Application US/09314268

i Patent No. 6346377

i GENERAL INFORMATION:

i APPLICANT: Doobdar, John
 i TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
 i FILE REFERENCE: 379/60902
 i CURRENT APPLICATION NUMBER: US/09/314,268
 i EARLIER APPLICATION NUMBER: 1999-03-19
 i EARLIER FILING DATE: 1999-05-18
 i NUMBER OF SEQ ID NOS: 179
 i SEQ ID NO: 35
 i SOFTWARE: PatentIn Ver. 2.1
 i LENGTH: 8
 i TYPE: PRT
 i ORGANISM: Artificial sequence
 i FEATURE:
 i OTHER INFORMATION: Description of Artificial Sequence: synthetic
 i OTHER INFORMATION: octapeptide antigen
 US-09-314-268-35

Query Match 66.7%; Score 50; DB 4; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PSPWAPKK 10
 |:|||||
 Db 1 PSPWAPKK 8

RESULT 15
 US-09-921-099A-21
 i Sequence 21; Application US/09921099A
 i Patent No. 6602707
 i GENERAL INFORMATION:
 i APPLICANT: Hefeneider, Steven
 i APPLICANT: Merkins, Louise
 i APPLICANT: Bennett, Robert
 i APPLICANT: Seiss, Donald
 i TITLE OF INVENTION: Mammalian DNA Binding Membrane-Associated Protein-encoding Gene a
 i FILE REFERENCE: 00-617-A
 i CURRENT APPLICATION NUMBER: US/09/921,099A
 i CURRENT FILING DATE: 2001-08-01
 i NUMBER OF SEQ ID NOS: 21
 i SOFTWARE: PatentIn version 3.0
 i SEQ ID NO: 21
 i LENGTH: 325
 i TYPE: PRT
 i ORGANISM: Saccharomyces cerevisiae
 US-09-921-099A-21

Query Match 65.3%; Score 49; DB 4; Length 325;
 Best Local Similarity 70.0%; Pred. No. 10;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 PKPSPWAPKK 10
 |:|||||
 Db 79 PRPSPWLPSK 88

Search completed: May 27, 2004, 16:22:37
 Job time : 9.69231 secs

This Page Blank (uspto)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:21:04 ; Search time 25.3846 Seconds
(without alignments)
132.009 Million cell updates/sec

Title: US-10-008-524A-168
Perfect score: 75
Sequence: 1 PKPSWPKHR 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 100%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
 1: /cgm2_6/ptodata/2/pubpaas/us07_pubcomb.pep:
 2: /cgm2_6/ptodata/2/pubpaas/us06_pubcomb.pep:
 3: /cgm2_6/ptodata/2/pubpaas/us05_pubcomb.pep:
 4: /cgm2_6/ptodata/2/pubpaas/us04_pubcomb.pep:
 5: /cgm2_6/ptodata/2/pubpaas/us03_pubcomb.pep:
 6: /cgm2_6/ptodata/2/pubpaas/us02_pubcomb.pep:
 7: /cgm2_6/ptodata/2/pubpaas/us01_pubcomb.pep:
 8: /cgm2_6/ptodata/2/pubpaas/us09a_pubcomb.pep:
 9: /cgm2_6/ptodata/2/pubpaas/us09b_pubcomb.pep:
 10: /cgm2_6/ptodata/2/pubpaas/us09c_pubcomb.pep:
 11: /cgm2_6/ptodata/2/pubpaas/us09d_pubcomb.pep:
 12: /cgm2_6/ptodata/2/pubpaas/us09_pubcomb.pep:
 13: /cgm2_6/ptodata/2/pubpaas/us1a_pubcomb.pep:
 14: /cgm2_6/ptodata/2/pubpaas/us1b_pubcomb.pep:
 15: /cgm2_6/ptodata/2/pubpaas/us1c_pubcomb.pep:
 16: /cgm2_6/ptodata/2/pubpaas/us10_pubcomb.pep:
 17: /cgm2_6/ptodata/2/pubpaas/us10_new_pub.pep:
 18: /cgm2_6/ptodata/2/pubpaas/us60_pubcomb.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	75	100.0	12	14	US-10-008-524A-168	Sequence 168, App
2	75	100.0	12	15	US-10-008-524A-168	Sequence 168, App
3	75	100.0	15	14	US-10-008-524A-167	Sequence 167, App
4	75	100.0	15	15	US-10-008-524A-167	Sequence 167, App
5	75	100.0	25	14	US-10-008-524A-4	Sequence 4, App
6	75	100.0	26	15	US-10-008-524A-4	Sequence 4, App
7	75	100.0	92	14	US-10-008-524A-90	Sequence 90, App
8	75	100.0	92	15	US-10-008-524A-90	Sequence 90, App
9	68	90.7	35	14	US-10-008-524A-117	Sequence 117, App
10	68	90.7	35	14	US-10-008-524A-117	Sequence 117, App
11	52	69.3	8	14	US-10-008-524A-33	Sequence 33, App
12	52	69.3	8	14	US-10-008-524A-37	Sequence 37, App
13	52	69.3	8	15	US-10-008-524A-33	Sequence 33, App
14	52	69.3	8	15	US-10-008-524A-37	Sequence 37, App
15	51	68.0	8	14	US-10-008-524A-36	Sequence 36, App

SEQUENCES

```

Sequence 36, App
Sequence 118, App
Sequence 118, App
Sequence 34, App
Sequence 35, App
Sequence 34, App
Sequence 35, App
Sequence 254044, App
Sequence 264398, App
Sequence 152, App
Sequence 32, App
Sequence 38, App
Sequence 32, App
Sequence 38, App
Sequence 119, App
Sequence 119, App
Sequence 71639, App
Sequence 233479, App
Sequence 30822, App
Sequence 125, App
Sequence 192, App
Sequence 2059, App
Sequence 3168, App
Sequence 6, App
Sequence 6, App
Sequence 910, App
Sequence 910, App
Sequence 910, App
Sequence 3041, App

```

ALIGNMENTS

```

RESULT 1
; Sequence 168, Application US/10008524A
; GENERAL INFORMATION:
; Publication No. US20030175682A1
; APPLICATION: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074A
; CURRENT APPLICATION NUMBER: US/10-008-524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 168
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
; US-10-008-524A-168

```

SUMMARY

```

Query Match 100.0%; Score 75; DB 14; Length 12;
Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;
Matches 12; Conservative 0;
Qy 1 PKPSWPKHR 12
Db 1 PKPSWPKHR 12

```

RESULTS

```

RESULT 2
; Sequence 168, Application US/10350719
; GENERAL INFORMATION:
; Publication No. US2003021972A1
; APPLICATION: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074A
; CURRENT APPLICATION NUMBER: US/10-008-524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 168
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
; US-10-008-524A-168

```

```

; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/350,719
; PRIORITY NUMBER: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIORITY NUMBER: GB0018140.4
; PRIOR FILING DATE: 2001-03-16
; PRIORITY NUMBER: GB0018140.4
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 168
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-168

RESULT 3
US-10-008-524A-167
Query Match 100.0%; Score 75; DB 15; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 1 PKPSPWPKHR 12

RESULT 4
US-10-350-719-167
Query Match 100.0%; Score 75; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 1 PKPSPWPKHR 15

RESULT 5
US-10-008-524A-4
Query Match 100.0%; Score 75; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 4 PKPSPWPKHR 15

RESULT 6
US-10-350-719-4
Query Match 100.0%; Score 75; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 4 PKPSPWPKHR 15

RESULT 7
US-10-350-719-4
Query Match 100.0%; Score 75; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0036;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 4 PKPSPWPKHR 15

RESULT 8
US-10-350-719-4
Query Match 100.0%; Score 75; DB 15; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PKPSPWPKHR 12
Db 4 PKPSPWPKHR 15

```

```

; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-117

Query Match 90.7%; Score 68; DB 14; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

RESULT 10
US-10-350-719-117
Query Match 90.7%; Score 68; DB 14; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 90
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-008-524A-90

Query Match 100.0%; Score 75; DB 14; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 1 PKPSPWAKKHR 12
Db 29 PKPSPWAKKHR 40

RESULT 8
US-10-350-719-90
Query Match 90.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: PCT/GB01/01176
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 117
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-350-719-117

Query Match 90.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

RESULT 11
US-10-008-524A-33
Query Match 90.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/10/008,524A
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-10-008-524A-117

Query Match 100.0%; Score 75; DB 15; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 1 PKPSPWAKKHR 12
Db 29 PKPSPWAKKHR 40

RESULT 9
US-10-008-524A-117
Query Match 90.7%; Score 68; DB 15; Length 35;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 2 KPSPWAKKHR 12
Db 1 KPSPWAKKHR 11

; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: US/10/008,524A
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-10-008-524A-117

Query Match 100.0%; Score 75; DB 15; Length 92;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy 1 PKPSPWAKKHR 12
Db 29 PKPSPWAKKHR 40

```

; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: Octapeptide antigen
US-10-008-524A-33

Query Match 69.3%; Score 52; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 1 PKPSPWAP 8
Db 1 PKPSPWAP 8

RESULT 12
US-10-008-524A-37
Sequence 37, Application US/10008524A
; Publication No. US20030175662A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/2162
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2003-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 37
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: octapeptide antigen
US-10-008-524A-37

Query Match 69.3%; Score 52; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 5 PWAPKKHR 12
Db 1 PWAPKKHR 8

RESULT 15
US-10-008-524A-36
Sequence 36, Application US/10008524A
; Publication No. US20030175662A1
; GENERAL INFORMATION:
; APPLICANT: Doorbar, John
; TITLE OF INVENTION: IMPROVEMENTS IN OR RELATING TO SCREENING FOR PAPILLOMA
; FILE REFERENCE: 18396/1074
; CURRENT APPLICATION NUMBER: US/10/008,524A
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 09/314,268
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 36
; LENGTH: 8
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: octapeptide antigen
US-10-008-524A-36

Query Match 69.3%; Score 52; DB 14; Length 8;
Best Local Similarity 100.0%; Pred. No. 1e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;

Qy 4 SPWAPKKH 11
Db 1 SPWAPKKH 11

Fri May 28 09:17:11 2004

us-10-008-524a-168.rapb

Page 5

Db 1 SPWAPKKH 8

Search completed: May 27, 2004, 16:34:19
Job time : 26.3846 secs

This Page Blank (up to)

A;Cross-references: GB:W74117
 C;Superfamily: Papillomavirus E4 protein
 C;Keywords: early protein

Query Match 77.3%; Score 58; DB 1; Length 96;
 Best Local Similarity 75.0%; Pred. No. 0.1;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PKPSPWPKXER 12
 Db 32 PKPAPWAPQKPR 43

RESULT 3
 W4WT31
 B4 protein - human papillomavirus type 31
 C;Species: human papillomavirus type 31
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 16-Jul-1999
 C;Accession: E32444
 R;Goldsbrough, M.D.; Disilvestre, D.; Temple, G.F.; Lorincz, A.T.
 Virology 171, 306-311, 1989
 A;Title: Nucleotide sequence of human papillomavirus type 31: a cervical neoplasia-associated
 A;Reference number: A94398; PMID:89299478; PMID:2545036
 A;Accession: E32444
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-102 <GOL>
 A;Cross-references: GB:J04353; NID:9333048; PIDN:AAA46949.1; PID:9459015
 A;Note: in Genbank entry PH31A the initiation codon UUG for residue 1 is translated as
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 69.3%; Score 52; DB 1; Length 102;
 Best Local Similarity 80.0%; Pred. No. 0.81;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PKPSPWPKXER 10
 Db 33 PKPAPWAPQKPR 42

RESULT 4
 AI0201
 beta-galactosidase (EC 3.2.1.23) [imported] - Yersinia pestis (strain CO92)
 C;Species: Yersinia pestis
 C;Accession: AI0201 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 R;Parthill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 Deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
 ill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
 Nature 413, 523-527, 2001
 A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A;Reference number: AB0001; MUID:11586360
 A;Accession: AI0201
 A;Status: Preliminary
 A;Molecule type: DNA
 A;Residues 1-1060 <KUR>
 A;Cross-references: GB:AL590842; PIDN:CAC90476.1; PID:915979691; GSPDB:GN00175
 C;Genetics:
 A;Gene: lacZ
 C;Superfamily: beta-galactosidase
 C;Keywords: glycosidase; hydrolase

Query Match 68.0%; Score 51; DB 2; Length 1060;
 Best Local Similarity 58.3%; Pred. No. 9.9;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PKPSPWPKXER 12
 Db 723 PKTPWPWSAQHR 734

RESULT 5
 S57977
 CCCH zinc finger protein CTH1 - yeast (Saccharomyces cerevisiae)
 N;Alternate name: protein YDB358.07c; protein YDR151C
 C;Species: Saccharomyces cerevisiae
 C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999
 C;Accession: S57977; JC5001
 R;Murphy, L.; Richards, C.; Harris, D.
 submitted to the EMBL Data Library, July 1995
 A;Reference number: S57977
 A;Accession: S57977
 A;Molecule type: DNA
 A;Residues: 1-325 <MUR>
 A;Cross-references: EMBL:Z50046; NID:9899393; PIDN:CAA90373.1; PID:9899400; MIPS:YDR151
 A;Experimental source: strain AB972
 R;Thompson, M.J.; Lai, W.S.; Taylor, G.A.; Blackshear, P.J.
 Gene 174, 225-233, 1996
 A;Title: Cloning and characterization of two yeast genes encoding members of the CCCH c
 A;Reference number: JC5001; MUID:37045817; PMID:8890739
 A;Accession: JC5001
 A;Molecule type: DNA
 A;Residues: 1-141, 'RV', 144-325 <THIO>
 A;Cross-references: GB:142133; NID:g1020082; PIDN:AAB39897.1; PID:91020083
 C;Comment: This protein belongs to the CCCH-type zinc finger protein family, and is a s
 C;Genetics:
 A;Gene: SGD:CTH1
 A;Cross-references: SGD:S0002558; MIPS:YDR151C
 A;Map position: 4R
 C;Keywords: zinc finger
 F;132-136:Region: nuclear location signal

Query Match 65.3%; Score 49; DB 2;
 Best Local Similarity 70.0%; Pred. No. 6.5;
 Matches 7; Conservative 1; Mismatches 2; Indels 0;
 Qy 1 PKPSPWPKXER 10
 Db 79 PRPSPWLPSK 88

RESULT 6
 T38817
 hypothetical protein SPAC4F10.13c - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Accession: T38817
 R;Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A;Reference number: Z21813
 A;Accession: T38817
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-992 <CON>
 A;Cross-references: EMBL:Z98980; NID:e1060691; PIDN:CAB11716.1; GSPDB:GN00066; SPDB:SPAD
 A;Experimental source: strain 972h-; cosmid c4F10
 C;Genetics:
 A;Gene: SPDB:SPAC4F10.13C
 A;Map position: 1
 A;Introns: 13/2

Query Match 64.0%; Score 48; DB 2;
 Best Local Similarity 69.2%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 2; Indels 2;
 Qy 1 PKPSPWPKXER 11
 Db 622 PRPSPWLPSK 634

RESULT 7
 A4004
 histidine decarboxylase (EC 4.1.1.22) - Enterobacter aerogenes
 C;Species: Enterobacter aerogenes

C;Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 18-Jun-1999
 C;Accession: A40004
 R;Kamath, A.V.; Vaaler, G.L.; Snell, E.E.
 J. Biol. Chem. 266, 9432-437, 1991
 A;Title: Pyridoxal phosphate-dependent histidine decarboxylases. Cloning, sequencing, and enzymes.
 A;Reference number: A40004; MUID:91236707; PMID:2033044
 A;Accession: A40004
 A;Status: not compared with conceptual translation
 A;Molecule type: DNA
 A;Residues: 1-378 <XAM>
 A;Cross-references: GB:M62745; NID:9435593; PIDN:AAA24802.1; PID:9435594
 C;Superfamily: Klebsiella histidine decarboxylase
 C;Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate P;233/Binding site: Pyridoxal phosphate (Lys) (covalent) #status predicted
 Query Match 62.7%; Score 47; DB 1; Length 378;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPKKH 11
 Db 329 PRPSENWVKKH 339

RESULT 8
 S72714 Lepb1170_F2_64 protein - Mycobacterium leprae
 C;Species: Mycobacterium leprae
 C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
 C;Accession: S72714
 R;Smith, D.R.; Robison, K.
 submitted to the EMBL Data Library, November 1993
 A;Description: Mycobacterium leprae cosmid B1170.
 A;Accession number: S72693
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-230 <SMI>
 C;Cross-references: EMBL:U00010; NID:9466780; PIDN:AA17078.1; PID:g466802
 C;Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1517

Query Match 61.3%; Score 46; DB 2; Length 230;
 Best Local Similarity 77.8%; Pred. No. 13;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPK 9
 Db 110 PKPSKWAPR 118

RESULT 9
 W4N151 E4 protein - human papillomavirus type 51
 C;Species: human papillomavirus type 51
 A;Note: host Homo sapiens (man)
 C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 27-Jan-1995
 C;Accession: C40415
 R;Lang, O.; Crum, C.P.; Silverstein, S.J.
 J. Virol. 65, 4216-4225, 1991
 A;Title: Biologic properties and nucleotide sequence analysis of human papillomavirus type 51
 A;Reference number: A40415; MUID:91303675; PMID:1649326
 A;Accession: C40415
 A;Status: translation not shown
 A;Molecule type: DNA
 A;Residues: 1-87 <LUN>
 C;Cross-references: GB:M62877
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 60.0%; Score 45; DB 1; Length 87;
 Best Local Similarity 66.7%; Pred. No. 7.3%;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 10
 T03301 rab3 effector protein Rim - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
 R;Wang, Y.; Okamoto, M.; Schmitz, F.; Hofmann, K.; Sudhof, T.C.
 Nature 388, 593-598, 1997
 A;Title: Rim is a putative Rab3 effector in regulating synaptic-vesicle fusion.
 A;Reference number: Z14897; MUID:97394473; PMID:9252191
 A;Accession: T03301
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-1553 <WAN>
 A;Cross-references: EMBL:AF007836; NID:92317777; PIDN:AAB65703.1; PMID:92317778
 A;Experimental source: tissue-type brain
 C;Genetics:
 C;Function:
 C;Description: Rim protein is proposed as Rab3-dependent regulator of synaptic-vesicle C;Keywords: GTP binding; zinc finger

Query Match 60.0%; Score 45; DB 2; Length 1553;
 Best Local Similarity 54.5%; Pred. No. 1.1e-02;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPKH 11
 Db 892 PQPSPBMPRRH 902

RESULT 11
 W4W118 E4 protein - human papillomavirus type 18
 C;Species: human papillomavirus type 18
 C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999
 C;Accession: E26251
 R;Cole, S.T.; Danos, O.
 J. Mol. Biol. 193, 599-608, 1987
 A;Title: Nucleotide sequence and comparative analysis of the human papillomavirus type
 A;Reference number: A92937; MUID:87283882; PMID:3039146
 A;Accession: E26251
 A;Molecule type: DNA
 A;Residues: 1-88 <COL>
 A;Cross-references: GB:X05015; NID:g60975; PIDN:CAA28668.1; PMID:g60980
 C;Superfamily: papillomavirus E4 protein
 C;Keywords: early protein

Query Match 58.7%; Score 44; DB 1; Length 88;
 Best Local Similarity 60.0%; Pred. No. 10;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 PKPSPWAPK 10
 Db 29 PAPCPWAPR 38

RESULT 12
 C64613 conserved hypothetical protein HP0747 - Helicobacter pylori (strain 26695)
 C;Species: Helicobacter pylori
 C;Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
 C;Accession: C64613
 R;Pomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKesson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, C.; Fujii, C.; Watthey, L.
 Nature 388, 539-547, 1997

A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A;Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A;Reference number: A64540; PMID:97394467; PMID:3252185
A;Accession: C64613
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-393 <POM>
A;Cross-references: GB:AE000587; GB:AE000511; NID:92313869; PID:AAD07796.1; PID:g231387

Query Match 57.3%; Score 43; DB 2; Length 393;
Best Local Similarity 70.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 PSPWAPKHR 12
Db 197 PVPNEEKHR 206

RESULT 13
D71900 hypothetical protein jhp0684 - Helicobacter pylori (strain J99)
C;Species: Helicobacter pylori
A;Variety: strain J99
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999
C;Accession: D71900
R;Alm, R.A.; Lang, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.J.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A;Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A;Reference number: A71800; MUID:99120557; PMID:9923682
A;Accession: D71900
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-400 <AN>
A;Cross-references: GB:AE001500; GB:AE001439; NID:94155238; PID:9415525
A;Experimental source: strain J99
C;Genetics:
A;Gene: jhp0684

Query Match 57.3%; Score 43; DB 2; Length 400;
Best Local Similarity 70.0%; Pred. No. 58;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 3 PSPWAPKHR 12
Db 197 PVPNEEKHR 206

RESULT 14
A31318 Glucose transporter-like protein - human
C;Species: Homo sapiens (man)
C;Date: 28-Feb-1999 #sequence_revision 28-Feb-1999 #text_change 24-Sep-1999
C;Accession: A31318
R;Fukumoto, H.; Seino, S.; Imura, H.; Seino, Y.; Eddy, R.L.; Fukushima, Y.; Byers, M.G.;
Proc. Natl. Acad. Sci. U.S.A. 95, 5434-5438, 1998
A;Title: Sequence, tissue distribution, and chromosomal localization of mRNA encoding a
A;Reference number: A31318; MUID:88289735; PMID:3399500
A;Accession: A31318
A;Molecule type: mRNA
A;Residues: 1-524 <FRK>
A;Cross-references: GB:J03810; NID:9187133; PID:AAA59514.1; PID:g307125
C;Genetics:

Query Match 57.3%; Score 43; DB 2; Length 524;
Best Local Similarity 60.0%; Pred. No. 75;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein

Scoring table:	BLOSUM62	Alignments
Gapped:	10.0	Gapped 0.5
Searched:	141681 seqs.	52070155 residues
Total number of hits satisfying chosen parameters:	141681	
Minimum DB seq length:	0	
Maximum DB seq length:	2000000000	
Post-processing:	Minimum Match 0%	
	Maximum Match 100%	
Database :	SwissProt_42::*	Listing first 45 summaries
Pred. No.	Score	Query Match Length DB ID Description
1	75	100.0 1 VE4 HPV16 P06222 human papill
2	58	77.3 95 1 VE4 HPV16 P27224 human Papil
3	52	69.3 102 1 VE4 HPV1 P17384 human Papil
4	49	65.3 325 1 CTHI YEAST P47366 saccharomyces
5	47	62.7 377 1 DCHS ENTATE P29877 enterobacte
6	45	60.0 87 1 VE4 HPV51 P26548 human papili
7	45	60.0 461 1 Y514 HUMAN P06069 homo sapien
8	45	60.0 1615 1 RIM1 RAT Q9jir4 rattus norv
9	45	60.0 1692 1 RIM1 HUMAN Q86rs5 homo sapien
10	44	58.7 88 1 VE4 HPV18 P06791 human papil
11	44	58.7 174 1 R122 SECCE P06336 secale cereale
12	43	57.3 393 1 TRMB HELPY P25443 helicobacte
13	43	57.3 400 1 TRMB HELP P092996 helicobacte
14	43	57.3 524 1 GTR2 HUMAN P11168 homo sapien
15	43	57.3 910 1 CHSA AMPQU P12564 amelomycetes
16	42	56.0 1822 1 ITB4 HUMAN P16144 homo sapien
17	41	54.7 232 1 TRMB AGRT5 P8uid4 agrobacteri
18	41	54.7 243 1 MYR2 SINAL P29738 sinapis alba
19	41	54.7 247 1 TRMB RHILIO P09813 rhizobium l
20	41	54.7 286 1 TRMB BIFLIO P08g74 bifidobacte
21	41	54.7 435 1 FUT9_ARATH Q9x177 arabidopsis
22	41	54.7 471 1 YHJ6 YEAST P38770 saccharomyces
23	41	54.7 670 1 ZN16 HUMAN P17020 homo sapien
24	41	54.7 778 1 SYQ_DROME P9y105 drosophila
25	41	54.7 1236 1 POLS_WEEV P13897 western egypt
26	41	54.7 1268 1 VGLN_HUMAN Q00341 homo sapien
27	41	54.7 1270 1 VGLN_CHICK P81021 gallus gallus
28	40	53.3 163 1 CU69 HUMAN P58556 homo sapien
29	40	53.3 202 1 VF56 MYCTU P10774 mycobacteri
30	40	53.3 282 1 SP0J_BACSU P26497 bacillus su
31	40	53.3 437 1 CAH9 MOUSE Q8vb5 mus musculus
32	40	53.3 478 1 BM3B_HUMAN P55107 homo sapien
33	40	53.3 542 1 ZYX_CHICK Q04584 gallus gallus

DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable E4 protein.
 OS Human papillomavirus type 35.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TAXID=10587;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92124753; PubMed=1310198;
 RA Marich J.B., Pontsler A.Y., Rice S.M., McGraw K.A., Dubensky T.W.;
 RT "The phylogenetic relationship and complete nucleotide sequence of
 human papillomavirus type 35.";
 RT Virology 186:770-776 (1992).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; M74117; AAA446365.2;
 DR PIR; C40824; W4W135.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 KW Early Protein.
 SQ SEQUENCE 96 AA; 10597 MW; AE452441BCD26F7C CRC64;
 SEQUENCE 96 AA; Score 58; DB 1; Length 96;
 Best Local Similarity 77.3%; Pred. No. 0.063;
 Matches 9; Conservative 75.0%;
 2; Mismatches 1; Indels 0; Gaps 0;
 DR 1 PKPSPWPKKR 12
 32 PKPAPNAPQKFR 43

Query Match 1 PKPSPWPKKR 12
 Best Local Similarity 77.3%; Pred. No. 0.063;
 Matches 9; Conservative 75.0%;
 2; Mismatches 1; Indels 0; Gaps 0;

Query Match 32 PKPAPNAPQKFR 43
 Best Local Similarity 75.0%; Pred. No. 0.063;
 Matches 9; Conservative 75.0%;
 2; Mismatches 1; Indels 0; Gaps 0;

RESULT 3
 VE4_HPV31 STANDARD PRT; 102 AA.
 AC P17304;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 01-AUG-1990 (Rel. 15, Last annotation update)
 DE Probable E4 protein.
 OS Human papillomavirus type 31.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TAXID=10585;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89399478; PubMed=2545036;
 RA Goldsborough M.D., Disilvestre D., Temple G.F., Lorincz A.T.;
 RT "Nucleotide sequence of human papillomavirus type 31: a cervical
 neoplasia-associated virus.";
 RT Virology 171:306-311 (1989).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; J04353; AAA446949.1; -.
 DR PIR; B32444; W4W131.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 KW Early protein.
 SQ SEQUENCE 102 AA; 11284 MW; 04E3C9E8ABC5CAAC CRC64;

Query Match 1 PKPSPWPKKR 12
 Best Local Similarity 70.0%; Pred. No. 3.7%;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Query Match 32 PKPAPNAPQKFR 43
 Best Local Similarity 70.0%; Pred. No. 3.7%;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Query Match 1 PKPSPWPKKR 12
 Best Local Similarity 65.3%; Pred. No. 3.7%;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Query Match 32 PKPAPNAPQKFR 43
 Best Local Similarity 65.3%; Pred. No. 3.7%;
 Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
 DCHS_ENTAE STANDARD PRT; 377 AA.
 ID DCHS_ENTAE
 AC P28577;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)

DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE Histidine decarboxylase (EC 4.1.1.22) (HDC).
GN
OS Enterobacter aerogenes (Aerobacter aerogenes); Enterobacteriales;
OC Bacteria; Proteobacteria; Gammaproteobacteria;
OT Enterobacteriaceae; Enterobacter.
OX NCBI_TaxID=548;
RN [1] -
RP SEQUENCE FROM N.A.
RX MEDLINE=9123670; PubMed=2033044;
RA Kamath A.V.; Vaaler G.L.; Snell E.B.;
RT "Pyridoxal-dependent histidine decarboxylases. Cloning, sequencing, and expression of genes from Klebsiella planticola and Enterobacter aerogenes and properties of the overexpressed enzymes.";
RL J. Biol. Chem. 266:9432-9437(1991).
CC -|- CATALYTIC ACTIVITY: L-histidine + CO(2).
CC -|- COFACTOR: Pyridoxal phosphate.
CC -|- SUBUNIT: Homotetramer (By similarity).
CC -|- SIMILARITY: Belongs to the group II decarboxylase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR M62745; AAA24302.1; -.
DR PIR: A40004; A40004
DR HAMAP; MF_0609; -; 1.
DR PROSITE; PS00282; PYRIDOXAL deC.
DR Pfam: PF00282; PYRIDOXAL deC; T.
DR PROSITE; PS00392; DDC_GAD_HDC_YDC; 1.
KW Lyase; Decarboxylases; Pyridoxal phosphate.
CC INIT MET 0 BY SIMILARITY.
FT BINDING 232 PYRIDOXAL PHOSPHATE (POTENTIAL).
SEQUENCE 377 AA; 42303 MW; 4C7A3334ACA7D6AE CRC64;
Query Match 62.7%; Score 47; DB 1; Length 377;
Best Local Similarity 72.7%; Pred. No. 8.1%;
Matches 8; Conservative 0; N mismatches 3; Indels 0; Gaps 0;

Qy 1 PKPSWPKKRR 11
Db 32B PKPSEWWKKH 338

RESULT 6
ID VEA4_HPV51 STANDARD: PRT; 87 AA.
AC P26548;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Probable E4 protein.
OS Human papillomavirus type 51.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
NCBI_TaxID=10595;
RN [1] -
RP SEQUENCE FROM N.A.
RX MEDLINE=1303675; PubMed=1649326;
RA Lungu O; Crum C.P.; Silverstein S.J.;
RT "Biological properties and nucleotide sequence analysis of human papillomavirus type 51";
RL Virol. 65:4216-4225(1991).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; M62877; -; NOT_ANNOTATED_CDS.
DR PIR; C40415; WAWL51.
DR InterPro; IPR003861; Papilloma_E4.
DR Pfam; PF02711; Pap_E4; 1.
KW Early protein.
SEQUENCE 87 AA; 9941 MW; 5F3DC38F86BF3990 CRC64;
Query Match 60.0%; Score 45; DB 1; Length 87;
Best Local Similarity 66.7%; Pred. No. 4;
Matches 8; Conservative 0; N mismatches 4; Indels 0; Gaps 0;
Qy 1 PKPSWPKKRR 12
Db 27 PLPPAWPKKPF 38

RESULT 7
ID Y514_HUMAN STANDARD: PRT; 461 AA.
AC 060259;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein KIA0514.
GN KIA0514.
OS Homo sapiens (Human).
OC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9666;
RN SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=9829045; PubMed=9628581;
RA Nagase T.; Ishikawa K.-I.; Miyajima N.; Tanaka A.; Kotani H., Nomura N.; Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";
RT RT
RN [2]
RN DNA Res. 5:31-39(1998).
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.; Schuler G.D., Klausner R.D.; Collins F.S.; Wagner L.; Shemesh C.M.; Schuler G.D., Altchul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K., Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Hsieh F., Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E., Brownstein M.J.; Usdin T.B.; Yoshiyuki S.; Garnier J.P.; Prange C., Raha S.S.; Loqueland J.; Peters G.D.; Mulahay S.J., Bosak S.A.; McEwan P.J.; McKernan K.J.; Malek J.A.; Gunaratne P.H., Richards S.; Worley K.C.; Hale S.; Garcia A.M.; Gay L.J.; Hulyk S.W., Villalon D.K.; Muzny D.M.; Sodergren E.J.; Lu X.; Gibbs R.A., Fahy J.; Helton E.; Ketteman M.; Madan A.; Rodriguez S.; Sanchez A., Whiting M.; Madan A.; Young A.C.; Shevchenko Y.; Bouffard G.G., Blakesley R.W.; Touchman J.W.; Schmutz J.; Myers R.M., Rodriguez A.C.; Grimwood J.; Green R.D.; Dickson M.C., Butterfield Y.S.N.; Krzywinski M.I.; Skalska U.; Smilus D.E., Scherzer A.; Schein J.E.; Jones S.J.M.; Marra M.A.; RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:1689-1693 (2002).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	or send an email to license@isb-sib.ch).
CC	EMBL; AB011086; BAA5440.1; -.
DR	EMBL; BC011672; AAH1672.1; -.
KW	Hypothetical protein.
SQ	SEQUENCE 461 AA; 47657 MW; C800B3C4F56BCBBD CRC64;
Query Match	Score 60.0%; Score 45; DB 1; Length 461;
Best Local Similarity	58.3%; Pred. No 19;
Matches	7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY	1 PKPSPWAPKKHR 12
DB	6 PEPGPWAPLSPR 17
RESULT 8	RIM1_RAT
ID	RIM1_RAT STANDARD; PRT; 1615 AA.
AC	Q9JIR8; O35168;
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 42, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Regulating synaptic membrane exocytosis protein 1 (Rab3-interacting molecule 1) (RIM 1).
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Rodentia; Murinae; Rattus.
OX	NCBI_TAXID=10116;
RN	[1]
SEQUENCE FROM N.A. (ISOFORM 2), AND INTERACTION WITH RAB3A AND RAB3C.	TISSUE-Brain;
RP	MEDLINE=97294473; PubMed=20252191;
RA	Wang Y., Okamoto M., Schmitz F., Hofmann K., Suedhof T.C.; Wang Y., Sugita S., Suedhof T.C.; "The RIM/NIM family of neuronal C2 domain proteins. Interactions with Rab3 and a new class of Src homology 3 domain proteins.";
RL	J. Biol. Chem. 275:20033-20044(2000).
CC	-!- FUNCTION: Rab effector involved in exocytosis. May act as scaffold protein.
CC	-!- SRBUNNT: Binds SNAP25, SYT1 and CACNA1B. Interaction with SYT1 is enhanced by calcium ions. Interaction with SNAP25 is weaker in the presence of calcium ions. Binds RAB3A, RAB3B and RAB3D that have been activated by GTP-binding. Binds UNC13 (By similarity).
CC	-!- SUBCELLULAR LOCATION: Peripheral membrane protein associated with plasma membranes from synaptic junctions. Not detected in synaptic vesicles. Detected in presynaptic nerve terminals close to the active zone. Detected in synaptic ribbons of ribbon synapses of retinal photoreceptor cells.
CC	-!- ALTERNATIVE PRODUCTS:
CC	Event=Alternative splicing; Named isoforms=2;
CC	Name=1; Synonyms=Rim1B;
CC	IsoId-Q9JIR4-1; Sequence=Displayed;
CC	Name=2;
CC	IsoId-Q9JIR4-2; Sequence=VSP 008172;
CC	-!- TISSUE SPECIFICITY: Highly expressed in hippocampus, brain cortex, cerebellum and olfactory bulb. Detected at lower levels in midbrain, hindbrain and spinal cord. Detected retina and in spinal cord motor neurons.
CC	-!- SIMILARITY: Contains 2 C2 domains.
CC	-!- SIMILARITY: Contains 1 FYVE-type zinc finger.
CC	-!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC	-!- SIMILARITY: Contains 1 Rab-binding (RabBD) domain.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
DR	EMBL; AF007836; AABE6703.1; -.
DR	BMBL; AF199333; AAFFB1655.1; -.
DR	PIR; T03301; T03301.
DR	HSSP; P21707; IRSY.
DR	IntezPro; IPR000008; C2.
DR	IntezPro; IPR008973; C2_CALB.
DR	InterPro; IPR001478; PDZ.
DR	InterPro; IPR003315; RPHA_effector.
DR	InterPro; IPR000306; Znf_FYVE.
DR	Pfam; PF00168; C2.
DR	Pfam; PF00595; PDZ.
DR	SMART; SM00228; PDZ.
DR	PROSITE; PS50004; C2_DOMAIN_2.
DR	PROSITE; PS50105; PDZ.
DR	PROSITE; PS50916; RABBD.
DR	PROSITE; PS50178; ZF_FYVE.
DR	KW Metal-binding; Zinc_Repeat; Zinc-Finger; Alternative splicing.
FT	DOMAIN 22 205 RAB-BINDING.
FT	ZNF_FING 193 FYVE-TYPE.
FT	DOMAIN 422 455 ALA-RICH.
FT	DOMAIN 444 483 PRO-RICH.
FT	DOMAIN 619 705 PDZ.
FT	DOMAIN 758 864 C2 DOMAIN 1.
FT	DOMAIN 1259 1301 C2 DOMAIN 2.
FT	DOMAIN 1461 1563 Missing (in isoform 2).
FT	VARSPLIC 1107 1168 Missing (in isoform 2).
FT	/FTID=vsp 008172.
SQ	SEQUENCE 1615 AA; 179654 MW; 80E76F74BF35FB7E CRC64;
Query Match	60.0%; Score 45; DB 1; Length 1615;
Best Local Similarity	54.5%; Pred. No. 60;
Matches	6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY	1 PKPSPWAPKKH 11
DB	: : ::
DB	892 PQPSPPMPRRH 902
RESULT 9	RIM1_HUMAN
ID	RIM1_HUMAN STANDARD; PRT; 1692 AA.
AC	Q9HBA5; Q9HBA6; Q9HBA7; Q9HBA8; Q9HBA9; Q9HBA10; Q9HBA11; Q9HBA12; Q9HBA13; Q9HBA14; Q9HBA15; Q9HBA16; Q9HBA17; Q9HBA18; Q9HBA19; Q9HBA20; Q9HBA21; Q9HBA22; Q9HBA23; Q9HBA24;
AC	Q9HBA5; Q9HBA6;
DT	10-OCT-2003 (Rel. 42, Created)
DT	10-OCT-2003 (Rel. 43, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	Regulating synaptic membrane exocytosis protein 1 (Rab3-interacting molecule 1) (RIM 1).
GN	RIM1 OR RIM1 OR KIAA0340.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Primates; Catarrhini; Hominoidea; Homo.
NCBI_TaxID	9606;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 1), ALTERNATIVE SPlicing, AND VARIANT
RP	CORD7 HIS-820.
RX	MEDLINE=22546706; PubMed=12659814;
RA	Johnson S., Halford S., Morris A.G., Patel R.J., Wilkie S.E., Hardcastle A.J., Moore A.T., Zhang K., Hunt D.M.;
RA	Genomic organisation and alternative splicing of human RIM1, a gene implicated in autosomal dominant cone-rod dystrophy (CORD7). ; Genomics 81:304-314 (2003).
RN	[2]

Query Match Score 45; DB 1; Length 1692;
 Best Local Similarity 54.5%; Pred. No. 62;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 PKPSPWPKH 11
 Db 878 PQPSPPMPRRH 888

RESULT 10
 VE4_HPV18 STANDARD PRT; 88 AA.
 AC P06791;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 01-JAN-1988 (Rel. 06, Last annotation update)
 DE Probable E4 protein.
 OS Human Papillomavirus type 18.
 OC dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 CX NCBI_TAXID=10562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87283822; PubMed=3-039146;
 RA Cole S.T.; Danos O.;
 RT "Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products.";
 RI J. Mol. Biol. 193:593-608 (1987).
 CC This SWISS-PROT entry is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

DR EMBL; X06515; CAA28668.1; -.
 DR PIR; B26251; WW118.
 DR InterPro; IPR003861; Papilloma_E4.
 DR Pfam; PF02711; Pap_E4; 1.
 DR Early protein.
 SQ 88 AA; 9857 MW; D05F6200BF367B10 CRC64;
 Query Match Score 44; DB 1; Length 88;
 Best Local Similarity 60.0%; Pred. No. 5.7%;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PKPSPWPKK 10
 Db 29 PAPCPWPKK 38

RESULT 11
 RL122 SECCE STANDARD PRT; 174 AA.
 JD R122 SECCE STANDARD PRT; 174 AA.
 AC P06036;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE 50S ribosomal protein L12-2, chloroplast precursor (CL12-2).
 GN RL12-2.
 OC Secale cereale (Rye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Secale; Poaceae; Pooidae.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv; Halo; TISSUE=Leaf;
 RD NCBI_TAXID=4550;

RA Schmidt M.W.; Pichl L.; Lepper M.; Feierabend J.;
 RT "Identification of the nuclear-encoded chloroplast ribosomal protein L12 of the monocotyledonous plant Secale cereale and sequencing of two different cDNAs with strong codon bias";
 RT Biochim. Biophys. Acta 1172:349-352 (1993).
 CC -!- SIMILARITY: Belongs to the L12P family of ribosomal proteins.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC EMBL; X63340; CAA48114.1; -.
 DR HSMP; P02392; 1C7E.
 DR InterPro; IPR00206; Ribosomal_L12.
 DR InterPro; IPR00206; Ribosomal_L12.
 DR Pfam; PF00542; Ribosomal_L12; T.
 DR ProDom; PD001346; Ribosomal_L12; 1.
 DR TIGRFAMS; TIGR00855; L12; 1.
 KW Ribosomal protein; Chloroplast; Transit peptide; Multigene family.
 FT CHLOROPLAST (POTENTIAL)
 FT TRANSIT 1 45 CHLOROPLAST (POTENTIAL)
 FT CHAIN 46 174 50S RIBOSOMAL PROTEIN L12-2.
 FT DOMAIN 16 24 PRO(SER-RICH)
 FT DOMAIN 82 96 ALA-RICH.
 SQ SEQUENCE 174 AA; 18319 MW; AAE20457356288C0 CRC64;
 Query Match Score 44; DB 1; Length 174;
 Best Local Similarity 66.7%; Pred. No. 11;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 PKPSPWPK 9
 Db 20 PSPPPWAPR 28

RESULT 12
 TRMB_HELPPY STANDARD PRT; 393 AA.
 ID TRMB_HELPPY STANDARD PRT; 393 AA.
 AC O25443;
 AC 025443;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DB DE Guanine-N(7)-methyltransferase (EC 2.1.1.33) (rRNA (m7G46) - methyltransferase).
 GN HP0747.
 OS Helicobacter pylori (Campylobacter; Epsilonproteobacteria; Campylobacteriales; Helicobacteraceae; Helicobacter).
 OC Hayes W.S.; Borodovsky M.; Karp P.D.; Smith H.O.; Fraser C.M.;
 OC Ventre J.C.;
 RN [1] NCBI_TAXID=210;
 RN SEQUENCE FROM N.A.
 RP STRAIN=26655 / ATCC 700392;
 RX MEDLINE=97194467; PubMed=2522185;
 RX MEDLINE=97194467; PubMed=97194467;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G., Fleischmann R.D., Ketchum K.A., Kline R.P., Gill S., Dougherty B.A., Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S., Lofthouse B., Richardson D., Dodson R., Khalsa H.G., Glodek A., McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K., Berg D.B., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M., Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M., Ventre J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter pylori.";
 RL Nature 388:539-547 (1997).
 CC -!- FUNCTION: Catalyzes the formation of N(7)-methylguanine at position 46 (m7G46) in tRNA (By Similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-homocysteine + tRNA containing N(7)-methylguanine.
 CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. Trmb

family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; MF_01057; -; 1.
 CC DR InterPro; IPR003358; Methyltransf_4.
 CC DR InterPro; IPR000051; SAM_bind.
 CC DR Pfam; PF02390; Methyltransf_4; 1.
 CC DR TIGRFAMS; TIGR00091; TIGR00091; 1.
 CC DR Transferase; Methyltransferase; tRNA processing; Complete proteome.
 CC DR SEQUENCE 400 AA; 46402 MW; 8B92B929;83D3573 CRC64;
 CC DR PIR; C64613; CS4613.
 CC DR TIGR; HP0747; -.
 CC DR InterPro; IPR004395; Cons_hypoth91.
 CC DR InterPro; IPR003158; Methyltransf_4.
 CC DR Pfam; PF02390; Methyltransf_4; 1.
 CC DR TIGRFAMS; TIGR00091; TIGR00091; 1.
 CC DR Transferase; Methyltransferase; tRNA processing; Complete proteome.
 CC DR SEQUENCE 393 AA; 45774 MW; C37EF02333F71D1 CRC64;
 SQ DR Query Match 57.3%; Score 43; DB 1; Length 393;
 DR Best Local Similarity 70.0%; Pred. No. 31;
 DR Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 DR Qy 3 PSPWAKKHR 12
 DR Db 197 PVPWNNEKKHR 206

RESULT 13
 TRMB HELPJ STANDARD; PRT; 400 AA.
 ID AC Q9ZJ56; 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE tRNA (guanine-N(7)-) -methyltransferase (EC 2.1.1.33) (tRNA(m7G46) - methyltransferase).
 DE JHP0684.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; BspIionproteobacteria; Campylobacterales;
 OC Helicobacterae; Helicobacter.
 NCBI TaxID=85463;

RN DR SEQUENCE FROM N.A.; MEDLINE=99120557; PubMed=9923682;
 RX RA Smith R.A., Ling L.-S.L., Moir D.T., King B.L., Brown B.D., Doig P.C.,
 RA Tummino P.J., Noonan B., Guidi B.C., de Jonge B.L., Carmel G.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RL Nature 397:176-180(1999)."
 CC -!- FUNCTION: Catalyzes the formation of N(7)-methylguanine at
 CC position 46 (m7G46) in tRNA (By similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing N(7)-methylguanine.
 CC -!- SIMILARITY: Belongs to the methyltransferase superfamily. Trmb
 CC family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; AE001500; AAD06272.1; -.

CC DR PIR; D71900; D71900;

DR DR HAMAP; MF_01057; -; 1.
 DR DR InterPro; IPR003358; Methyltransf_4.
 DR DR InterPro; IPR000051; SAM_bind.
 DR DR Pfam; PF02390; Methyltransf_4; 1.
 DR DR TIGRFAMS; TIGR00091; TIGR00091; 1.
 DR DR Transferase; Methyltransferase; tRNA processing; Complete proteome.
 DR DR SEQUENCE 400 AA; 46402 MW; 8B92B929;83D3573 CRC64;
 DR DR PIR; A31318; A31318.
 DR DR Genew; HGNC:110006; SLC2A2.
 DR DR GK; P11168; -.
 DR DR NM_138160; -.
 DR DR GO; GO:0005687; C:integral to plasma membrane; TAS.
 DR DR GO; GO:0005624; C:membrane fraction; TAS.
 DR DR GO; GO:0005355; P:Glucose transporter activity; TAS.
 DR DR GO; GO:0005975; P:carbohydrate metabolism; TAS.
 DR DR GO; GO:0015758; P:Glucose transport; TAS.
 DR DR InterPro; IPR007114; MFS.

DR InterPro; IPRO005828; Sub transporter.
 DR InterPro; IPRO005829; Sug transporter.
 DR InterPro; IPRO03663; Sugar transporter.
 PFam; PF00083; sugar_tr_1.
 PRINTS; PRO0171; SUGTRANSPORT.
 DR TIGRFAMS; TIGR00879; SP_1.
 DR PROSITE; PS00216; SUGAR TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR TRANSPORT_2; 1.
 KW Transmembrane; Sugar transport; Glycoprotein;
 MultiGene family; Polymorphism.

FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 11 31 1 (POTENTIAL).
 FT TRANSEM 32 98 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 99 119 2 (POTENTIAL).
 FT DOMAIN 120 127 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 128 148 3 (POTENTIAL).
 FT DOMAIN 149 158 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 159 179 4 (POTENTIAL).
 FT DOMAIN 180 187 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 188 208 5 (POTENTIAL).
 FT DOMAIN 209 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 218 238 6 (POTENTIAL).
 FT DOMAIN 239 303 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 304 324 7 (POTENTIAL).
 FT DOMAIN 325 338 EXTRACELLULAR (POTENTIAL).
 FT TRANSEM 339 359 8 (POTENTIAL).
 FT DOMAIN 360 368 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 369 389 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 390 400 9 (POTENTIAL).
 FT TRANSEM 401 421 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 422 433 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 434 454 10 (POTENTIAL).
 FT DOMAIN 455 461 CYTOPLASMIC (POTENTIAL).
 FT TRANSEM 462 482 11 (POTENTIAL).
 FT DOMAIN 483 524 EXTRACELLULAR (POTENTIAL).
 FT CARBOYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 68 68 P -> L
 FT FTId=VAR 007169 /FTId=VAR 007169.
 FT VARIANT 101 101 V -> I (in dbSNP 18003572).
 FT VARIANT 110 110 T -> I (in dbSNP 5400).
 FT VARIANT 478 478 L -> V (in dbSNP 5397).
 FT FTId=VAR 014720 /FTId=VAR 014720.
 SQ SEQUENCE 524 AA; 57489 MW; DA600577207EC033 CRC64;

Query Match 57.3%; Score 43; DB 1; Length 524;
 Best Local Similarity 60.0%; Pred. No. 41;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 PKRSPWPKR 10
 Db 7 PSRSPWAPQR 18

RESULT 15
 CHSA_AMPOU AMPQU STANDARD; PRT; 910 AA.
 AC Q12564;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chitin synthase A (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl transferase A) (Class-I chitin synthase A).
 GN CHSA.
 OS Ampelomyces quisqualis.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Leptosphaeriales; mitosporic Leptosphaeraceae;
 OC Ampelomyces.
 OX NCBI_TaxID=50730;
 RN [1] SEQUENCE FROM N.A.

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:12:58 ; search time 23.0769 Seconds
(164.069 Million cell updates/sec)

Title: US-10-008-524A-168
Perfect score: 75
Sequence: 1 PKSPWAPKHR 12

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTRIMBL 25:
 1: sp_archaea:
 2: sp_bacteria:
 3: sp_fungi:
 4: sp_human:
 5: sp_invertebrate:
 6: sp_mammal:
 7: sp_mhc:
 8: sp_organelle:
 9: sp_phage:
 10: sp_plant:
 11: sp_rabbit:
 12: sp_virus:
 13: sp_vertebrate:
 14: sp_unclassified:
 15: sp_rvirus:
 16: sp_bacteriophage:
 17: sp_archeap:
 18: sp_bacteria:
 19: sp_fungi:
 20: sp_mammal:
 21: sp_mhc:
 22: sp_organelle:
 23: sp_phage:
 24: sp_rabbit:
 25: sp_vertebrate:
 26: sp_virus:
 27: sp_bacteriophage:
 28: sp_archeap:
 29: sp_bacteria:
 30: sp_fungi:
 31: sp_mammal:
 32: sp_mhc:
 33: sp_organelle:
 34: sp_phage:
 35: sp_rabbit:
 36: sp_vertebrate:
 37: sp_virus:
 38: sp_bacteriophage:
 39: sp_archeap:
 40: sp_bacteria:
 41: sp_fungi:
 42: sp_mammal:
 43: sp_mhc:
 44: sp_organelle:
 45: sp_phage:

ALIGNMENTS

Result No.	Score	Query	Match	Length	DB ID	Description
1	75	100.0	95	12	Q918t2 human papil	Q918t2 human papil
2	75	100.0	95	12	Q918u0 human papil	Q918u0 human papil
3	75	100.0	95	12	Q918t7 human papil	Q918t7 human papil
4	75	100.0	95	12	Q8BSP2 human papil	Q8BSP2 human papil
5	75	100.0	95	12	Q8BSN7 human papil	Q8BSN7 human papil
6	75	100.0	95	12	Q80MM0 human papil	Q80MM0 human papil
7	52	69.3	88	12	Q913Z6 human papil	Q913Z6 human papil
8	52	69.3	467	10	Q9A3T7 oryza sativ	Q9A3T7 oryza sativ
9	52	69.3	467	10	Q7XC25 oryza sativ	Q7XC25 oryza sativ
10	51	68.0	464	12	Q91DM2 cuparia herb	Q91DM2 cuparia herb
11	51	68.0	1060	16	Q8Zfp0 Yersinia pe	Q8Zfp0 Yersinia pe
12	49	65.3	2527	5	Q95W83 Plasmodium	Q95W83 Plasmodium
13	48	64.0	992	3	O36025 schizosacch	O36025 schizosacch
14	48	64.0	1242	12	Q9Pw8 eastern equ	Q9Pw8 eastern equ
15	48	64.0	1242	12	Q9Pw7 eastern equ	Q9Pw7 eastern equ
16	48	64.0	1242	12	Q9Pw6 eastern equ	Q9Pw6 eastern equ

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARES

Result	No.	Score	Query	Match	Length	DB ID	Description
RESULT 1	[1]	Q918t2	ID	Q918t2	PRELIMINARY;	PRT;	95 AA.
RC		AC	Q918t2;				
RA		DT	01-DBC-2001 (TREMBLrel.	19,	Created)		
RT		DT	01-MAR-2001 (TREMBLrel.	19,	Last sequence update)		
RL		DT	01-MAR-2002 (TREMBLrel.	20,	Last annotation update)		
DR		DB	E4 Protein (Fragment).				
GN		GN	E4.				
OS		OS	Human papillomavirus type 16.				
OC		OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;				
RN		RN	Papillomavirus.				
[1]			NCBI_TaxID=10581;				
SEQUENCE FROM N.A.							
RP							
SEQUENCE FROM N.A.							
RC			STRAIN=IPV1622CC7;				
RC			NON_TER 1 1				
RA			FT PFP02711; Pap_E4_1.				
RA			SQ SEQUENCE 95 AA; 10342 MW;				
RT			"Sequence variation and physical state of human Papillomavirus type 16 cervical cancer isolates from Australia and New Caledonia.";				
RL			Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.				
DR			InterPro: IPR003861; AAL014051; -.				
DR			PFam: PF02711; Pap_E4_1.				
FT			NON_TER 1 1				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				
FT			Best Local Similarity 100.0%; Mismatches 0; Indels 0; Gaps 0;				
FT			Score 75; DB 12; Length 95;				

OC Schizosaccharomyces.
 RN [1] NCBI_TaxID:4896;
 RP SEQUENCE FROM N.A.
 RC STRAIN=N9721-;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Wood V.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z98910; CAB1176.1; -.
 DR PTRI: T38817; T38817.
 DR GenBank; SPONBIE; SPACAF10.13c; -.
 DR InterPro; IPR003169; GYF.
 DR Pfam; PF02123; GYF; 1.
 DR SMART; SM00444; GYF; 1.
 DR Prosite; PS50839; GYF; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 992 AA; 106112 MW; 4751B990B0C39345 CRC64;
 Query Match 64.0%; Score 48; DB 3; Length 992;
 Best Local Similarity 69.2%; Pred. No. 35;
 Matches 9; Conservative 0; Mismatches 2; Indels 2; Gaps 1;

QY 1 PKPSPW-APKXH 11
 Db 622 PKPSPWPKSLPPKX 634

RESULT 14
 Q9PZWB PRELIMINARY; PRT; 1242 AA.
 AC Q9PZWB; PRELIMINARY; PRT; 1242 AA.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-OCT-2003 (TREMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Structural polyprotein.
 OS Eastern equine encephalitis virus (Eastern equine encephalomyelitis virus).
 OC ssRNA positive-strand viruses, no DNA stage; Togaviridae;
 OC Alphavirus.
 OX NCBI_TaxID:11021;
 RN [1] TAXID=11021;
 RP SEQUENCE FROM N.A.
 RC STRAIN=PA6-435731;
 RA Brault A.C., Powers A.M., Kang W., Tesh R.B., Shope R.E., Weaver S.C.;
 RT "Genetic and Antigenic Diversity among Eastern Equine encephalitis viruses from North, Central and South America.";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF159560; AA04802.1; -.
 DR HSSP; P03315; IVCP.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0019028; C:viral capsid; IEA.
 DR GO; GO:0004252; C:viral envelope; IEA.
 DR GO; GO:0005198; F:serine-type endopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR002548; Alpha_E1_glycop.
 DR InterPro; IPR000936; Alpha_E2_glycop.
 DR InterPro; IPR02533; Alpha_B3_glycop.
 DR InterPro; IPR009003; Cys_Ser_Trypsin.
 DR InterPro; IPR000336; Flavi_glycoprote.
 DR InterPro; IPR07110; Ig-like.
 DR InterPro; IPR00930; Peptidase_S3.
 DR Pfam; PF00944; Alpha_core; 1.
 DR PF01589; Alpha_E1_glycop; 1.
 DR Pfam; PF00943; Alpha_E2_glycop; 1.
 DR Pfam; PF01563; Alpha_E3_glycop; 1.
 DR PRINTS; PR00798; TOGAVIRIN.
 KW Polyprotein.
 SQ SEQUENCE 1242 AA; 137344 MW; 2759F20F690B4A13 CRC64;

Query Match 64.0%; Score 48; DB 12; Length 1242;
 Best Local Similarity 75.0%; Pred. No. 44;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 PKPSPWAPKKHR 12
 Db 73 PKPSPWAPKKHR 84

Search completed: May 27, 2004, 16:20:57
 Job time : 24.0769 secs

SEQUENCE 1242 AA; 137436 MW; 25A0990E69D16636 CRC64;
 Query Match 64.0%; Score 48; DB 12; Length 1242;
 Best Local Similarity 75.0%; Pred. No. 44;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

This Page Blank (up to)

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2004, 16:03:03 ; Search time: 69.2308 Seconds
(without alignments)
102.031 Million cell updates/sec

Title: US-10-008-524A-4
Perfect score: 142
Sequence: 1 RPIPKPSEWAPKXHRRRLSDQDSQT 25

Scoring table: BLCSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:
Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing First 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseq_ip1980s:*
- 2: geneseq_ip1990s:*
- 3: geneseq_ip2000s:*
- 4: geneseq_ip2001s:*
- 5: geneseq_ip2002s:*
- 6: geneseq_ip2003as:*
- 7: geneseq_ip2003bs:*
- 8: geneseq_ip2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	131.5	92.6	26	2	AAW62277		Aaw62277 HPV16 hyd
2	92.6	26	5	AAU75260	Hydrophil		Aar15260 Hydropophil
3	110.5	77.8	30	2	AAR15564		Aar15564 Immunopep
4	98.5	69.4	20	2	AAR15562		Aar15562 Immunopep
5	82	57.7	15	2	AAW62278		Aaw62278 HPV16 hyd
6	81	57.0	14	2	AAR14288		Aar14288 Seroreact
7	75	52.8	12	2	AAN62279		Aaw62279 HPV16 hyd
8	66	46.5	11	2	AAR14287		Aar14287 Seroreact
9	61.5	43.3	1053	7	ADE54638		Ade54638 Human Pro
10	61.5	43.3	1553	7	ADE54636		Ade54636 Rat Prote
11	61	43.0	51	4	AAU51469		Aau51469 Propionib
12	61	43.0	51	6	ABM47988		Abm47988 Propionib
13	58	40.8	985	4	ABBS59214		Abbs59214 Drosophil
14	57	40.1	230	5	ABU05501		Abu05501 M. tuberculosis
15	56	39.4	105	4	AAO03023		Aao03023 Human pol
16	56	39.4	394	4	ABG16932		Abg16932 Novel hum
17	55	39.1	525	5	AB47787		Ab47787 Protein e
18	54	38.0	77	4	AAU55993		Aau55993 Propionib
19	54	38.0	77	6	ABM52512		Abm52512 Propionib
20	54	38.0	121	4	ABG19332		Abg19332 Novel hum
21	54	38.0	177	7	ADM74128		Adm74128 DNA clon
22	54	38.0	682	7	ADC37582		Aoc37582 Human nuc
23	54	38.0	683	5	ABB97217		Abb97217 Novel hum
24	54	38.0	2451	4	ABB71574		Abb71574 Drosophil
25	53.5	37.7	598	2	AAM99065		Aao15355 Human sig

RESULT 1
AAW62277
ID AAW62277 standard; peptide; 26 AA.
XX
AC
XX
DT 24-SEP-1998 (first entry)
XX
DE HPV16 hydrophilic region peptide #1.
XX
KW HPV16; human papilloma virus; epithelial tumour; cervical cancer;
KW preancerous cervical lesion; screening; detection; infection; cervix;
KW HPV E4.
XX
OS Human papillomavirus.
XX
PN WO925145-A1.
XX
PD 11-JUN-1998.
XX
PP 03-DEC-1997; 97TWO-GB003321.
XX
PR 03-DEC-1996; 96GBB-00025142.
PR 05-SEP-1997; 97GBB-00018745.
XX
(MED1-) MEDICAL RES COUNCIL.
XX
PA Doorbar J;
XX
PI DR WPI; 1998-333497/29.
XX
PT Detecting papilloma virus infection using molecule binding to E4 protein
PT - useful, e.g. in screening for pre-cancerous cervical lesions and to
PT determine type(s) of human papilloma virus infecting human patients.
XX
PS Claim 9; Page 37; 52pp; English.
XX
CC A new method has been developed for detecting a papilloma virus infection
CC in an organism. The method comprises: (i) obtaining a sample of cells
CC from the potential infection site; (ii) contacting the cells with a
CC molecule binding specifically to papilloma virus E4 protein, and (iii)
CC monitoring the binding. The method is useful to detect papilloma virus
CC infections in organisms (especially mammals) and especially HPV
CC infections (e.g. with HPV16, 18, 31, 35, 45, 51, 56, 58 or 61) in humans.
CC Papilloma viruses cause epithelial tumours in humans varying in severity
CC depending on the infection site and HPV type involved. The method is
CC particularly useful to determine Papilloma infection in the mammalian
CC cervix and especially to screen for pre-cancerous cervical lesions in

humans, since over 90% of cervical carcinoma patients show cervical HPV infection. It is also useful to determine the type(s) of HPV infection in a patient, by using a molecule binding specifically to a subset of HPV E4 proteins. This is important, since progression to malignant disease (and hence clinical prognosis) is dependent on HPV type. Molecules capable of binding E4 are also useful to target anticancer/antiviral agents capable of destroying papilloma viruses and/or papilloma virus-infected cells. The present sequence represents a specifically claimed HPV16 peptide found in the hydrophilic region.

Sequence 26 AA;
 SQ Query Match Score 131.5; DB 2; Length 26;
 Best Local Similarity 96.2%; Pred. No. 6e-11;
 Matches 25; Conservative 0; Mismatches 0;
 Db 1 RPIPKPSWAPKKGHLRL-SDQDSQTP 25
 1 RPIPKPSWAPKKGHLSSDQDSQTP 26

RESULT 2

AAU75260

standard; peptide; 26 AA.

XX

AAU75260;

AC

AAU75260;

XX

DT

21-MAY-2002 (first entry)

XX

DE

Hydrophilic region of HPV16 E4 protein.

XX

conventional papilloma smear testing programmes. Compared with previous methods of screening, the new method has reduced chances of false negatives occurring, requires fewer samples to gain the same amount of information, and alleviates the need for repeated or further testing. The present sequence representing the hydrophilic region of the HPV16 E4 protein binds antibody molecules.

XX

Sequence 26 AA;

SQ Query Match Score 131.5; DB 5; Length 26;

Best Local Similarity 92.6%; Pred. No. 6e-11;

Matches 25; Conservative 0; Mismatches 0;

Db 1 RPIPKPSWAPKKGHLRL-SDQDSQTP 25

1 RPIPKPSWAPKKGHLSSDQDSQTP 26

RESULT 3

ID AAR15564 standard; protein; 30 AA.

XX

AC AAR15564;

XX

DT 02-MAR-1992 (first entry)

XX

DE Immunopeptide #4 derived from HPV16 E4 peptide.

XX

XX cervical cancer; cervical intraepithelial neoplasia; CIN; wart;

KW squamous cell carcinoma; ELISA; HPV 16.

XX Synthetic.

OS PN WO9118294 A.

XX PR 11-MAY-1990; 90SE-00001705.

XX PA (MEDS-) MEDSCAND AB.

XX PI Dillner J, Dillner L, Cheng HM;

XX XX DR WPI: 1991-369396/50.

XX PT Diagnosis of human Papilloma virus infection and PV-carrying tumours -

PT using synthetic peptide(s) to detect virus specific antigen-antibody

PT complexes by immunoassay.

XX Disclosure; Page 38; 72pp; English.

XX This is one of a large number of peptides which have been synthesised on the basis of the amino acid sequences for the E2, E4, E7, L1 or L2 proteins of HPV 1, 5, 6, 8, 11, 16, 18, 31 and 33. The selection of peptide sequences was based on the assumption that an immunoreactive region might be situated in the same relative region of a protein from different HPV types. The peptides were used in diagnostic immunoassays to detect HPV-infection. See AAR15523-R15601

SQ Sequence 30 AA;

Query Match Score 110.5; DB 2; Length 30;

Best Local Similarity 97.8%; Pred. No. 5.2e-08;

Matches 21; Conservative 0; Mismatches 0;

Indels 1; Gaps 1;

OY 1 RPIPKPSWAPKKGHLRL-SDQD 21

7 RPIPKPSWAPKKGHLSSDQD 28

DB

RESULT 4.

PI Muller M, Gissmann L;
 XX WPI; 1991-304643/42.
 XX Sero-active epitope(s) of human papilloma-virus 16 proteins - for use
 PT vaccines and diagnosis.
 XX
 PS Claim 1; Page 11; 15pp; English.
 XX This is one of two seroreactive epitopes identified from HPV16 protein
 CC E4. Peptides containing this epitope are also claimed and can be used to
 CC generate antibodies to HPV. See AARI4287-R14302 and AAQ14168-Q14171.
 XX Sequence 14 AA:
 SQ 5 KPSPWAKKHKRLS 18
 CC Query Match 57.0%; Score 81; DB 2; Length 14;
 CC Best Local Similarity 100.0%; Pred. No. 0.00026;
 CC Matches 14; Conservative 0; Mismatches 0;
 CC Indels 0; Gaps 0;
 CC DB 1 KPSPWAKKHKRLS 14
 CC
 RESULT 7
 ID AAW62279 standard; peptide; 12 AA.
 XX AAW62279;
 AC AAW62279;
 DT 24-SEP-1998 (first entry)
 XX DE HPV16 hydrophilic region peptide #3.
 XX KW HPV16; human papilloma virus; epithelial tumour; cervical cancer;
 KW precancerous cervical lesion; screening; detection; infection; cervix;
 KW HPV E4.
 OS Human papillomavirus.
 XX FN WO9825145-A1.
 XX PR 11-JUN-1998.
 DD 03-DEC-1997; 97WO-GB003321.
 XX PF 03-DEC-1996; 96GB-00025142.
 XX PR 05-SEP-1997; 97GB-00018145.
 PA (MEDI-) MEDICAL RES COUNCIL.
 XX PI Doorbar J;
 XX DR 1998-333497/29.
 XX Detecting papilloma virus infection using molecule binding to E4 protein
 PT - useful, e.g. in screening for pre-cancerous cervical lesions and to
 PT determine type(s) of human papilloma virus infecting human patients.
 XX PS Claim 11; Page 37; 52pp; English.
 XX A new method has been developed for detecting a papilloma virus infection
 CC in an organism. The method comprises: (i) obtaining a sample of cells
 CC from the potential infection site; (ii) contacting the cells with a
 CC molecule binding specifically to papilloma virus E4 protein, and (iii)
 CC monitoring the binding. The method is useful to detect Papilloma virus
 CC infections in organisms (e.g. with HPV16, 18, 33, 35, 45, 51, 56, 58 or 61) in humans.
 CC Papilloma viruses cause epithelial tumors in humans varying in severity
 CC depending on the infection site and HPV type involved. The method is
 CC particularly useful to determine papilloma infection in the mammalian
 CC cervix and especially to screen for pre-cancerous cervical lesions in
 CC humans, since over 90% of cervical carcinoma patients show cervical HPV

CC infection. It is also useful to determine the type(s) of HPV infection in
 CC a patient, by using a molecule binding specifically to a subset of HPV E4
 CC proteins. This is important, since progression to malignant disease (and
 CC hence clinical prognosis) is dependent on HPV type. Molecules capable of
 CC binding E4 are also useful to target anticancer/antiviral agents capable
 CC of destroying papilloma viruses and/or papilloma virus-infected cells.
 CC The present sequence represents a specifically claimed HPV16 peptide
 CC found in the hydrophilic region
 XX Sequence 12 AA;
 SQ Query Match 52.8%; Score 75; DB 2; Length 12;
 CC Best Local Similarity 100.0%; Pred. No. 0.0015;
 CC Matches 12; Conservative 0; Mismatches 0;
 CC Indels 0; Gaps 0;
 QY 4 PKPSPWAKKHKR 15
 CC DB 1 PKPSPWAKKHKR 12
 CC
 RESULT 8
 ID AARI4287 standard; peptide; 11 AA.
 XX AARI4287;
 AC AARI4287;
 DT 02-JAN-1992 (first entry)
 XX Seroreactive epitope #1 of HPV16 protein E4.
 DE XX
 XX HPV16-dependent human disease; E6; E7; Li.
 KW XX Synthetic.
 OS XX
 PN EP451550-A.
 XX PD 16-OCT-1991.
 XX PP 19-MAR-1991; 91EP-00104197.
 XX PR 20-MAR-1990; 90EP-00105222.
 XX PA (BEHRINGERWERKE AG.
 XX PI Muller M, Gissmann L;
 XX DR WPI; 1991-304643/42.
 XX PA Sero-active epitope(s) of human papilloma-virus 16 proteins - for use
 PT vaccines and diagnosis.
 XX PS Claim 1; Page 11; 15pp; English.
 XX This is one of two seroreactive epitopes identified from HPV16 protein
 CC B4. Peptides containing this epitope are also claimed and can be used to
 CC generate antibodies to HPV. See also AARI4288-R14302 and AAQ14168-Q14171.
 XX Sequence 11 AA;
 SQ Query Match 46.5%; Score 66; DB 2; Length 11;
 CC Best Local Similarity 100.0%; Pred. No. 0.023;
 CC Matches 11; Conservative 0; Mismatches 0;
 CC Indels 0; Gaps 0;
 QY 3 IPKPSWAPKK 13
 CC DB 1 IPKPSWAPKK 11
 CC
 RESULT 9
 AD54648
 ID AD54638 standard; protein; 1053 AA.
 XX AC AD54638;
 AC

XX 29-JAN-2004 (first entry)
 DT DE Human Protein BAA20798, SEQ ID NO 443.
 XX KW Human; pain; neuronal tissue; gene therapy;
 spinal segmental nerve injury; chronic constriction injury; CCI;
 spared nerve injury; SNI; Chung.
 XX OS Homo sapiens.
 PN WO2003016475-A2.
 XX PD 27-FEB-2003.
 XX PP 14-AUG-2002; 2002WO-US025765.
 XX PR 14-AUG-2001; 2001US-03121475.
 PR 01-NOV-2001; 2001US-0346382P.
 26-NOV-2001; 2001US-0333347P.
 XX PA (GEHO) GEN HOSPITAL CORP.
 (FARB) BAYER AG.
 XX PI Woolf C, D'urso D, Befort K, Costigan M;
 XX DR WPI: 2003-268312/26.
 DR GENBANK; BAA20798.
 XX PT New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.
 PS Claim 1; Page: 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence given in the
 CC specification, a method for identifying a compound which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition,
 CC a method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC the sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 Sequence 1053 AA;
 SQ Query Match 43.3%; Score 61.5; DB 7; Length 1053;
 Best Local Similarity 33.3%; Pred. No. 11; Mismatches 2; Indels 13; Gaps 1;
 Matches 11; Conservative 7; Mismatches 2; Indels 13; Gaps 1;
 DB 890 PIPKPSPWAPKHH-----RRLSDQD 21
 DB 890 PLPQESPMPRRHGESSSKLQRORSQRISDSD 922
 QY 2 PIPKPSPWAPKHH-----RRLSDQD 21
 QY 2 PIPKPSPWAPKHH-----RRLSDQD 21
 Query Match 43.3%; Score 61.5; DB 7; Length 1553;
 Best Local Similarity 33.3%; Pred. No. 17;
 Matches 11; Conservative 7; Mismatches 2; Indels 13; Gaps 1;
 Sequence 1553 AA;

Best Local Similarity	44.4%;	Pred.	No.	0.56;						DT	08-APR-2003	(first entry)
Matches	12; Conserved	4;	Mismatches	9;	Indels	2;	Gaps	1;		XX	M. tuberculosis and M. leprae marker protein #152.	
Qy	1 RPIPKPSAPKPK- <u>HRRLLSDQSOTP</u> 25									KW	Mycobacteriosis; survival; virulence; protective antigen; vaccine;	
DB	1 RPSFPPTRCPQQNSHRRPFGQELAP 27									KW	mycobacterial disease; tuberculosis; leprosy.	
RESULT 13										XX	Mycobacterium tuberculosis.	
ABB59214	ABB59214 standard; protein; 985 AA.									OS	Mycobacterium leprae.	
XX										XX	W0200274903-A2.	
AC										XX	26-SEP-2002.	
PD										XX	22-FEB-2002; 2002WO-IB001973.	
XX										XX	22-FEB-2001; 2001US-0270123P.	
XX										XX	(INSP) INST PASTEUR.	
DE										XX	PA	
XX										XX	PI Cole S;	
XX										XX	PI XX	
XX										XX	DR WPI; 2002-759885/82.	
CC										XX	PT PT	Identifying and selecting genes for survival or virulence of mycobacteria by a comparative genomic analysis of the sequences of Mycobacterium tuberculosis and M. leprae.
CC										XX	PT	
CC										XX	XX	
CC										XX	PS Claim 17; Page 322-323; 874pp; English.	
CC										XX	CC This invention relates to a novel method for identifying essential genes for survival or virulence of mycobacteria species. The method comprises aligning the genomic sequence of a first mycobacterium species on a genomic sequence of a second mycobacterium species and selecting a polynucleotide sequence that is highly conserved in both genomes with no counterparts in other bacterial genomic sequences and that corresponds to an essential gene for the survival or virulence of mycobacterium species. The method of the invention is useful for detecting M. tuberculosis or M. leprae infection. The method reduces the number of potential new targets and protective antigens for new drugs and vaccine compositions to treat and prevent mycobacterial diseases, particularly tuberculosis and leprosy. The present sequence represents a marker protein from Mycobacterium tuberculosis and Mycobacterium leprae identified using the method of the invention	
CC										XX	CC Sequence 230 AA;	
CC										XX	Query Match 40.1%; Score 57; DB 5; Length 230; Best Local Similarity 57.9%; Pred. No. 9.6; Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;	
CC										XX	QY 4 PKPSPWAPKKHRLLSDQDS 22	
CC										XX	Db 110 PKPSKWAPRFLARLINGDS 128	
CC										XX	RESULT 15	
CC										XX	AA003023 standard; protein; 105 AA.	
CC										XX	ID AA003023	
CC										XX	AA003023;	
CC										XX	DT 06-NOV-2001	(first entry)
CC										XX	DB Human polypeptide SEQ ID NO 16915.	
CC										XX	XX Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haemopoiesis;	
CC										XX	KW tissue growth factor; immunomodulator; cancer; leukaemia; nervous system disorders; arthritis; inflammation.	
CC										XX	KW Homo sapiens.	
CC										XX	ABU05501_14	
CC										XX	ABU05501 standard; protein; 230 AA.	
CC										XX	ABU05501;	

PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX PF 26-FEB-2001; 2001WO-US004927.
 XX PR 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX PA (HYSEQ INC.)
 PI HYSEQ INC.
 XX Tang YT, Liu C, Dirmanac RT;
 DR 2001-514838/56.
 DR N-PSDB; AA182954.
 XX PR Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PR and treating e.g. leukemia, inflammation and immune disorders.
 XX PS Claim 20; SEQ ID NO 16915; 1399pp + Sequence Listing; English.
 XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp://wipo.int/pub/published_pct_sequences
 XX SQ Sequence 105 AA:
 Query Match 39.4%; Score 56; DB 4; Length 105;
 Best Local Similarity 81.8%; Pred. No. 5.8;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 4 PKSPMAPKKH 14
 | | | | | | | |
 Db 82 PCSPWAKTH 92

Search completed: May 27, 2004, 16:18:27
 Job time : 72.2308 secs